MULTIVARIATE STATISTICAL PROCESS CONTROL APPROACH FOR PERFORMANCE EVALUATION IN CALL CENTERS

by
Tolga Ahmet Kalaycı
B.S., Industrial Engineering, İstanbul Technical University, 2008

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
The requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University
2011

ACKNOWLEDGEMENTS

I would like to thank my advisor Prof. Ali Rıza Kaylan for his thoughtful guidance and help during the preparation of this dissertation. Thanks are also due to the members of my committee, Prof. Mehmet C. Çamurdan and Assoc. Prof. Wolfgang Hörmann, for their helpful comments and suggestions.

I wish to express my appreciation to my parents for their support. They helped me to stay focused on my work by providing me with a happy and peaceful home.

I am also grateful to call center executives that I was in contact during my data collection period. I appreciate them for their assistance and collaboration.

Finally, special thanks to The Scientific and Technological Research Council of Turkey (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu - TÜBİTAK) which supported me financially towards my M.S. degree in Industrial Engineering.

ABSTRACT

MULTIVARIATE STATISTICAL PROCESS CONTROL APPROACH FOR PERFORMANCE EVALUATION IN CALL CENTERS

In this thesis, a multivariate performance assessment model for the inbound department of a call center was proposed. The model is treated with principal component analysis (PCA). The main reasons for the employment of PCA are its simplified signal decomposition ability, adaptability in recursive approaches and autocorrelated structure of the collected data. Due to the time-varying dynamics of the call center processes, conventional static PCA algorithms are not suitable for such an application. For this reason, a new recursive PCA algorithm which is capable of tracing time-varying nature of call center dynamics, was constructed as an alternative to the previously introduced recursive algorithms in the literature. In the proposed model, six performance indicators were defined for the inbound call center process. While these indicators represent the variables on the control charts, the days stand for observations. Consequently, a signaling point on any statistical control chart designates to a day at which the performance of the call center process has significantly drifted. Two years of data were processed through the model and the results were analyzed. Additionally, to make our performance assessment model easier to understand and more practical, a decision support table involving the mostly encountered types of signaling days and their practical interpretations was given at the end of the study. Also, the study showed that the PCA method which has been mostly used for industrial processes, may efficiently be adapted into a service system performance evaluation works.

ÖZET

ÇAĞRI MERKEZLERİNDE PERFORMANS DEĞERLENDİRMESİNE ÇOK DEĞİŞKENLİ İSTATİSTİKSEL SÜREÇ KONTROLÜ YAKLAŞIMI

Bu tezde, bir çağrı merkezinin gelen çağrı birimi için çok değişkenli bir performans değerlendirme modeli önerilmiştir. Model, temel bileşenler analizi (TBA) yöntemi üzerine kurulmuştur. TBA yönteminin kullanılmasının temel nedenleri, bu yöntemin sinyal ayrıştırmasındaki yalınlığı, özyineli yaklaşımlara uygunluğu ve üzerinde çalışılan veride özilinti olmasıdır. Çağrı merkezi süreçlerinin dinamik yapısı nedeniyle, geleneksel durağan TBA algoritmaları böyle bir çalışma için uygun değildir. Bu nedenle, önceki araştırmalarda önerilmiş olan diğer özyineli algoritmalara alternatif olabilecek, çağrı merkezi dinamiklerinin durağan olmayan yapısına ayak uydurabilen yeni bir özyineli TBA algoritması oluşturulmuştur. Önerilen modelde, bir çağrı merkezinin gelen çağrı süreci için altı performans göstergesi tanımlanmıştır. Bu göstergeler, kontrol şemalarındaki değişkenlere karşılık gelirken, günler ise gözlemlere karşılık gelmektedir. Dolayısıyla, kurulan model dâhilinde, istatistiksel kontrol şemalarında oluşan sinyaller, çağrı merkezinin süreç performansının anlamlı bir biçimde kaydığı günleri işaret etmektedir. İki senelik veri model kapsamında incelenmiş, alınan sonuçlar analiz edilmiştir. Ayrıca, geliştirdiğimiz performans değerlendirme modelini daha kolay anlaşılır ve uygulanabilir bir şekile sokmak için, en çok karşılaşılan sinyal veren gün tiplerini ve bunların yorumlarını içeren bir karar destek tablosu çalışmanın sonuna eklenmiştir. Bunların yanında, yaptığımız çalışma, şimdiye kadar çoğunlukla endüstriyel süreçlere uygulanmış TBA metodunun, etkin bir biçimde hizmet sistemi performans değerlendirme çalışmalarına uygulanabileceğinin bir göstergesi olmuştur.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT iv
ÖZET v
LIST OF FIGURES viii
LIST OF TABLES xii
LIST OF SYMBOLS xiii
LIST OF ACRONYMS/ABBREVIATIONS xv
1. INTRODUCTION 1
    1.1. Call Center Processes 4
2. LITERATURE REVIEW 7
    2.1. Performance Management in a Call Center 7
    2.2. Multivariate Statistical Process Control and Principal Component Analysis 10
    2.3. Performance Evaluation Applications 16
        2.3.1. SPC Techniques in Performance Evaluation 17
        2.3.2. Performance Evaluation Works for Call Centers 20
3. METHODOLOGY 25
    3.1. Process Overview 25
    3.2. Problem Definition 28
    3.3. Objectives and Model Development 30
        3.3.1. Definition of New Performance Metrics 31
    3.4. Data Collection and Analysis 34
4. PROPOSED ALGORITHM AND ITS APPLICATION 40
    4.1. Algorithm Description 45
    4.2. Monte Carlo Study of the Proposed Algorithm 49
    4.3. Parameter selection 52
        4.3.1. Parameter Selection for Weekdays 53
        4.3.2. Parameter Selection for Weekends 61
    4.4. Detailed Analysis of Selected Scenarios and Decision Support Table 65
5. CONCLUSIONS AND FUTURE WORKS 76

APPENDIX A: AUTOCORRELATION FUNCTIONS OF TWO-YEAR DATA SET 79
APPENDIX B: AUTOCORRELATION FUNCTIONS OF WEEKDAYS AND WEEKENDS (ROW DATA) 81
APPENDIX C: AUTOCORRELATION FUNCTIONS OF WEEKDAYS AND WEEKENDS (DESEASONALIZED DATA) 85
APPENDIX D: MATLAB COMMANDS FOR THE PERIODIC WEIGHTED MOVING AVERAGE PCA ALGORITHM 87
APPENDIX E: MATLAB COMMANDS FOR THE MONTE CARLO STUDY (FIRST POLICY) 92
APPENDIX F: MATLAB COMMANDS FOR THE MONTE CARLO STUDY (SECOND POLICY) 95
REFERENCES 98

LIST OF FIGURES

Figure 1.1. Origination and the development steps of the study. 3
Figure 1.2. A call center view. 4
Figure 1.3. Call handling system. 6
Figure 2.1. Call center performance indicators. 9
Figure 3.1. Interaction between a customer and a sale-group agent. 26
Figure 3.2. Current system illustration. 33
Figure 4.1. Steps of the proposed multivariate process performance evaluation method. 48
Figure 4.2. *SPE* chart for main parameters Φ(20,0.6) (Weekdays). 53
Figure 4.3. T^2 chart for main parameters Φ(20,0.6) (Weekdays). 54
Figure 4.4. *SPE* chart for main parameters Φ(40,0.6) (Weekdays). 55
Figure 4.5. T^2 chart for main parameters Φ(40,0.6) (Weekdays). 56
Figure 4.6. *SPE* chart for main parameters Φ(60,0.6) (Weekdays). 56
Figure 4.7. T^2 chart for main parameters Φ(60,0.6) (Weekdays). 57
Figure 4.8. *SPE* chart for main parameters Φ(20,0.7) (Weekdays). 57
Figure 4.9. T^2 chart for main parameters Φ(20,0.7) (Weekdays). 58
Figure 4.10. *SPE* chart for main parameters Φ(40,0.7) (Weekdays). 59
Figure 4.11. T^2 chart for main parameters Φ(40,0.7) (Weekdays). 59
Figure 4.12. *SPE* chart for main parameters Φ(60,0.7) (Weekdays). 60
Figure 4.13. T^2 chart for main parameters Φ(60,0.7) (Weekdays). 60
Figure 4.14. *SPE* chart for main parameters Φ(24,0.6) (Weekends). 61

Figure 4.15. T^2 chart for main parameters Φ(24,0.6) (Weekends). 62

Figure 4.16. *SPE* chart for main parameters Φ(32,0.6) (Weekends). 63

Figure 4.17. T^2 chart for main parameters Φ(32,0.6) (Weekends). 63

Figure 4.18. *SPE* chart for main parameters Φ(32,0.7) (Weekends). 64

Figure 4.19. T^2 chart for main parameters Φ(32,0.7) (Weekends). 64

Figure 4.20. Contributions for weekday 102. 66

Figure 4.21. Contributions for weekday 130. 67

Figure 4.22. Contributions for weekday 150. 68

Figure 4.23. Contributions for weekday 256. 68

Figure 4.24. Contributions for weekend 114. 69

Figure 4.25. Contributions for weekday 261. 69

Figure 4.26. Contributions for weekend 46. 70

Figure 4.27. Contributions for weekday 312. 70

Figure 4.28. Contributions for weekday 420. 71

Figure 4.29. Contributions for weekdays 403-406. 72

Figure 4.30. Proposed performance assessment system. 73

Figure A.1. Autocorrelation function for *AP*. 79

Figure A.2. Autocorrelation function for *ACD*. 79

Figure A.3. Autocorrelation function for *RT*. 79

Figure A.4. Autocorrelation function for *CpA*. 80

Figure A.5. Autocorrelation function for *SpC*. 80

Figure A.6. Autocorrelation function for *WT*. 80

Figure B.1. Original and first order difference autocorrelation functions for *AP* (weekdays). 81

Figure B.2. Original and first order difference autocorrelation functions for *ACD* (weekdays). 81

Figure B.3. Original and first order difference autocorrelation functions for *RT* (weekdays). 81

Figure B.4. Original and first order difference autocorrelation functions for *CpA* (weekdays). 82

Figure B.5. Original and first order difference autocorrelation functions for *SpC* (weekdays). 82

Figure B.6. Original and first order difference autocorrelation functions for *WT* (weekdays). 82

Figure B.7. Original and first order difference autocorrelation functions for *AP* (weekends). 83

Figure B.8. Original and first order difference autocorrelation functions for *ACD* (weekends). 83

Figure B.9. Original and first order difference autocorrelation functions for *RT* (weekends). 83

Figure B.10. Original and first order difference autocorrelation functions for *CpA* (weekends). 84

Figure B.11. Original and first order difference autocorrelation functions for *SpC* (weekends). 84

Figure B.12. Original and first order difference autocorrelation functions for *WT* (weekends). 84

Figure C.1. Original and first order difference autocorrelation functions for deseasonalized *CpA* (weekdays). 85

Figure C.2. Original and first order difference autocorrelation functions for deseasonalized *SpC* (weekdays). 85

Figure C.3. Original and first order difference autocorrelation functions for deseasonalized *WT* (weekdays). 85

Figure C.4. Original and first order difference autocorrelation functions for deseasonalized *ACD* (weekends). 86

Figure C.5. Original and first order difference autocorrelation functions for deseasonalized *CpA* (weekends). 86

Figure C.6. Original and first order difference autocorrelation functions for deseasonalized *WT* (weekends). 86

LIST OF TABLES

Table 3.1. Correlations for months 1-4. 35

Table 3.2. Correlations for months 5-16. 35

Table 3.3. Correlations for months 17-24. 36

Table 3.4. Eigenvalues and condition indices. 37

Table 3.5. Seasonal indices for weekdays. 38

Table 3.6. Seasonal indices for weekends. 39

Table 4.1. Results for three types of injected signals. 51

Table 4.2. Results of weekday model scenarios. 61

Table 4.3. Results of weekend model scenarios. 65

Table 4.4. Details of constructed models. 65

Table 4.5. Counterparts of performance metrics. 73

Table 4.6. Decision Support Table. 74

LIST OF SYMBOLS

C	Lower triangular matrix obtained by Cholesky decomposition of S
C_α	Normal deviate cutting of an area of α under the standard normal distribution
D	D statistic
E	Standardized residual matrix
e_i	i^{th} column of E matrix
F	Inverse of the F distribution
h_0	A function of θ_1, θ_2, and θ_3
K	Moving window half-width parameter
k	Number of principal components to be retained in the model
L	Diagonal matrix containing the first k eigenvalues of S as its diagonal entries
l_i	i^{th} eigenvalue of the covariance matrix
n	Number of observations
p	Number of variables
S	Covariance matrix
$\mathrm{S}_{\mathrm{MAD}}$	MAD estimate
SPE	Squared prediction error
$\bar{s}_t$	Weighted standard deviation vector
$\mathrm{S}_{\mathrm{X}_{\mathrm{generated}}}$	Covariance matrix of $\mathrm{X}_{\mathrm{generated}}$
t	Rejection threshold
T^2	Hotelling's multivariate statistic
U	Matrix containing the eigenvectors of S as its columns

U_k	Matrix containing the first k columns of U
X	Standardized data matrix
$X_{generated}$	Multinormal data matrix
X_t	Historical data set
$\bar{x}_t$	Weighted mean vector
x_t	New observation vector
$x^{\dagger}$	Median of the related data sequence
Y_t	Temporary data matrix
Z	*p x n* matrix composing of *n* independent standard normal variables
z	Principal component scores matrix
α	Significance level
μ	Mean matrix
σ_i	Standard deviation of the i^{th} variable
θ_1	Sum of the variances of unused principal components
θ_2	Sum of squares of the variances of unused principal components
θ_3	Sum of third powers of the variances of unused principal components

LIST OF ACRONYMS/ABBREVIATIONS

AB	Average Number of Agents Busy
ACD	Automatic Call Distributor
AP	Answer Percentage
ARIMA	Autoregressive Integrated Moving Average
CA	Number of calls answered
CpA	Call per Agent
CQ	Number of Calls Queued
CQT	Number of Calls in Queue at Time t
CTI	Computer Telephone Integration
CUSUM	Cumulative Sum
DT	Dominant Period
EWMA	Exponentially Weighted Moving Average
HDS	Historical Data Set
IC	Number of Incoming Calls
IVR	Interactive Voice Response
MAD	Median Absolute Deviation From the Median
MCUSUM	Multivariate Cumulative Sum
MEWMA	Multivariate Exponentially Weighted Moving Average
MSPC	Multivariate Statistical Process Control
MYT	Mason Young Tracy
PABX	Private Automatic Branch Exchange

PCA	Principal Component Analysis
PLS	Partial Least Squares
PVE	Percent Variance Explained
PWMA	Periodic Weighted Moving Average
SPC	Statistical Process Control
SpC	Sales per Call
SL	Service Level
WF	Weighting Factor

1. INTRODUCTION

In today's highly competitive markets, urgent need for continuous improvement makes an effective performance evaluation process an important necessity for all kind of industries. Call centers which serve as an integral part of almost all business fields are one of the leading industries requiring close inspection and continuous improvement.

The first introduction of the call center concept was at the end of 1960's as a demand and complaint conveyance platform. At these years, most companies in United States provided toll-free lines for their customers. At the beginning of 1970's, automatic call distributer (ACD) was first used by Continental Airlines. Following the first introduction of this concept, big companies began to turn their physical spaces in which inbound calls are handled into organized call centers. After gaining substantial cost-related benefits provided by this new technology integration and customer relationships, call centers have got one level upper in their development process. Together with enhanced business skills and growing capacities, call centers reached their current highly value-adding and integrated structure.

Although 30 years before, the call centers were considered as a competitive plus, they have turned out to be a competitive must in recent times. They are solidly integrated in most of business fields and as the competitiveness increases, economic part of their role is growing. Most enterprises which are in an intense communication process with their customers, involve a call center in their structure. Particularly, companies like banks airlines, telecommunication companies regard call centers as a crucial part of their businesses. These communication departments may be in an integrated form within the corporate structure or they may be outsourced. Both cases result in a rapidly growing call center industry. 1019 call centers recruiting approximately 50000 people most of which are young people are currently operating in Turkey (Economic Policy Research Foundation of Turkey, 2011). Despite negative economical developments happened in 2009, call centers market proved its growing potential. In 2010, Turkey call center market volume has increased to 1.2 billion TL with a 26% growth compared to previous year while this volume is \$ 3.5 billion in United States and \$ 350 billion worldwide (Call Centers Association, 2010).

Considering these economic and quality related roles of the call center industry, a successfully applied performance evaluation for call centers can make significant contributions to service quality and cost efficiency of the whole company. Due to dynamic and continuous nature of call center activities, an online performance monitoring which is able to adapt in this changing nature, is a must for a successful performance evaluation.

Performance monitoring is a process combining operational intelligence and monitoring tools to inform decision makers about significant performance changes. Multivariate statistical process control (MSPC) methods were proven to be effective tools in performance monitoring (Xiong *et al*., 2006) (Miletic *et al*., 2004).

This study proposes a multivariate performance assessment model for the inbound department of a call center. The model is treated with principal component analysis (PCA). The main reasons for the employment of PCA are its adaptability in nonstationary models, simple signal decomposition ability and autocorrelated structure of the collected data. Due to the time-varying dynamics of the call center processes, conventional static PCA algorithms are not suitable for such an application. For this reason, a new recursive PCA algorithm was constructed as an alternative to the previously introduced recursive algorithms in the literature. MATLAB software was used to construct and execute the algorithm.

In the proposed model, six performance indicators were defined for the inbound call center process. While these indicators represent the variables on the control charts, the days stand for observations. Consequently, a signaling point on any statistical control chart designates to a day at which the performance of the call center process has significantly drifted.

Additionally, to make our performance assessment model easier to understand and more practical, a decision support table involving the mostly encountered types of signaling days and their practical interpretations will be given at the end of the study.

As a helpful guideline to the reader about the flow of this thesis, the origination and the development steps of this study are given in Figure 1.1.

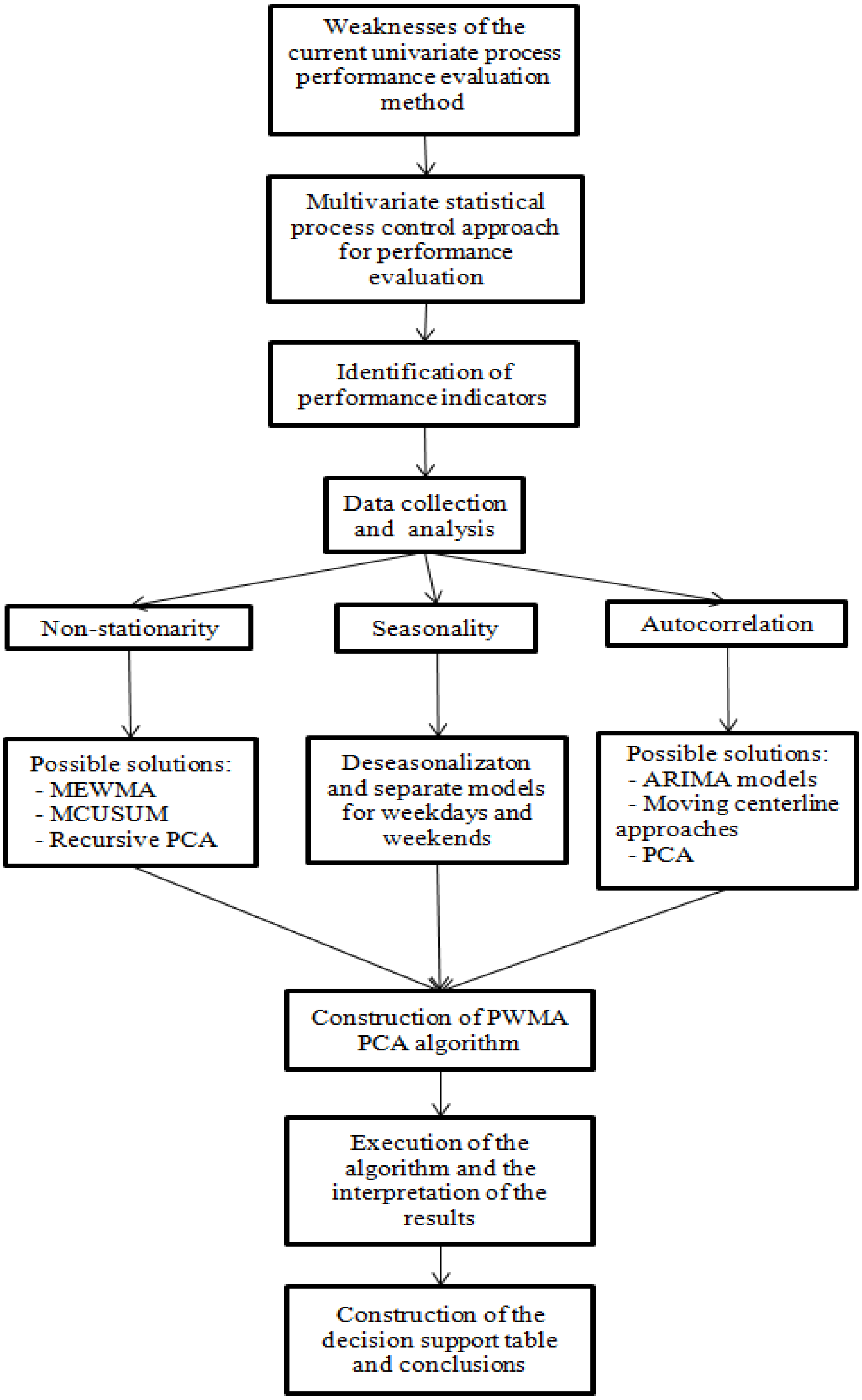

Figure 1.1. Origination and the development steps of the study.

1.1. Call Center Processes

A call center consists of various resources such as agents, computers and different kind of telecommunication devices. To envision a relatively big call center, we can think of a large room with separated open space agent desks on which is a computer.

.

Figure 1.2. A call center view.

Call centers may differ in terms of various criteria such as size, geographical dispersion, organization of work, handled call type. Apart from the organization of work which is about involving cross-trained or skill-based trained agents, the main characteristic of a call center is the type of call it handles; Inbound or outbound calls. As a call center may handle both types of calls which is the case it is called “blended call center”, it may be specialized on just one type. An inbound call means an incoming call which is created by someone outside the call center. Help desk or reservation calls may be given as examples of inbound calls. An outbound call means an outgoing call which is created from within the center such as telemarketing calls. Since our study is on inbound process performance of a call center, further details will focus on this type of services.

Main telecommunication equipments of an inbound call center are private automatic branch exchange (PABX), interactive voice response (IVR) unit, automatic call distributer (ACD) unit and computer telephone integration (CTI) middleware. Once an inbound call is initiated, it is connected through the PABX to the IVR unit which queries customers on their needs. During these queries in the IVR menu, in most of the routes that a caller may follow, he/she is asked for an identifying number such as card or customer number. At this point two cases may happen; Customer may get enough information via the voice response unit and hang up the phone or he/she may desire to speak with an agent. At the latter case, the customer stops interacting with IVR and he/she gets transferred to the ACD unit and joins in the queue. ACD is a sophisticated switch connected via PBAX to call center agents and is responsible for routing the call depending on predefined criteria. These criteria may be about callers preferred language, topic privileges or required skills. After joining in the queue, the call is routed to an appropriate available agent. In the case where all agents are busy, the call is kept "on-hold" until a feasible agent is available. In the inbound call center process which we will be working on, no skill-based routing is applied but all calls are joined is the same queue and assigned to first available agent depending on first in first out criteria. Stolen notice and credit card password topics are exceptions of this rule since they directly come to the first row because of their potential for a crucial case. It is possible to lose some of the calls during this waiting period due to reasons such as impatience, unacceptably long waiting times or technical problems. Callers who do not abandon the queue are eventually connected to an agent.

Once the interaction between the customer and the agent begins, callers are queried in the database depending on their identification number. The middleware performing this query is called computer telephone integration (CTI) technology which works as a customer relationship management tool. Consequently, the agent is provided with customer's personal records like history of customer's previous calls, service preferences or past sale figures. This system is called "screen pop-up system". It makes possible sale proposals display on agent's screen while eliminating additional times for the agent to ask for customer's personal data and improving standardization of service times (Gans *et al*., 2003).

Interaction between the customer and the agent may end in several ways; the customer may have a fully satisfactory service and hang up the phone or for a particular reason, he/she may be transferred to another official. In both cases, the agent spends some time for after call activities. Obviously the agent will keep being unavailable during this relatively short period. With the termination of after call working period, the agent becomes available to handle a new call. Below is the description of call handling system.

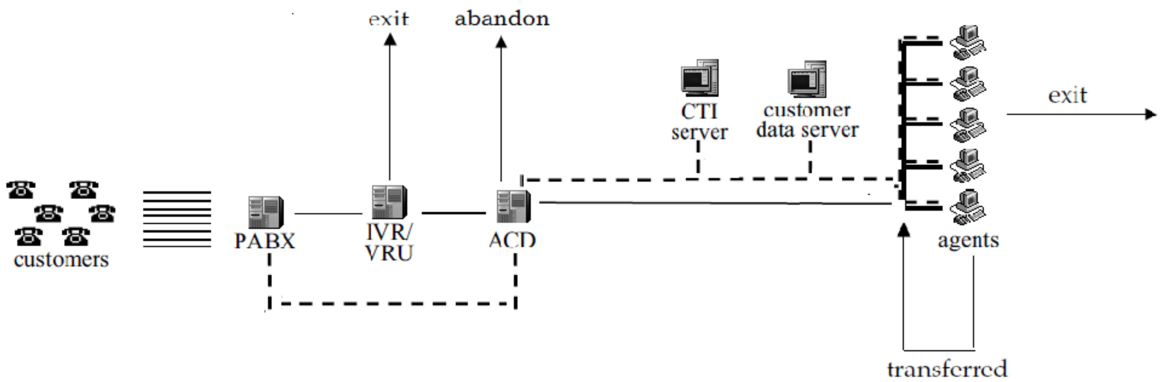

Figure 1.3. Call handling system.

Chapter 2 includes the literature review about performance management in call centers and gives the necessary theoretical background for PCA. Additionally, previously introduced call center performance evaluation works are extensively discussed. Chapter 3 introduces the methodology followed to develop the proposed model. Also, the answers of question such as “Why is this study required?” and “What is intended in this study?” are given in this part of the thesis. Description of the proposed algorithm and its application on a two-year data set composes the content of Chapter 4. Additionally, the decision support table covering practical signal interpretations is included in this part of the study. In Chapter 5, contributions and conclusions of the study are summarized and possible future works are briefly described.

2. LITERATURE REVIEW

2.1. Performance Management in a Call Center

Call center performance management is a quantitative management approach to improve service quality and efficiency of a call center. Its evaluation process is formed by performance metrics that are defined depending on corporation goals. The main aims of performance management may be stated as providing objective results about the performance of the call center and of its agents and ensuring improvement. As (Fluss, 2007) stated, with a successfully applied performance management, constructing a framework for matching the goals of a call center with those of the enterprise is a possible achievement in strategic level. Also, it was stated in the same paper that concerning the tactical level it is a good systematic approach to detect weaknesses and strengths by using objectives, performance indicators and evaluation techniques.

We can mainly talk about two types of performances in a call center; operational performance and financial performance. Although these two are closely related to each other, as the result of challenging economic conditions emphasizing cost containment and revenue maximization, financial performance for call centers began to differentiate from operational performance in recent times. Since the main interest of our research is in operational part, no more details will be given on financial aspects. Two subgroups of performance evaluations are discussed in the literature about the operational level. First one is process performance measurement which is a general performance evaluation for the call handling process. The other one is agent efficiency measurement which is an individual performance evaluation. Various general and individual performance indicators are used in the literature. These indicators may be summarized in two categories. The first category comprises of indicators that are related to call center statistics.

- *Number of calls in queue at time t (CQT)*: Number of customers who have completed interacting with IVR and who are currently waiting to connect to an agent via ACD.

- *Number of incoming calls (IC)*: Total number of calls initiated outside the call center and connected to IVR system via PABX.

- *Number of calls queued (CQ): T*otal number of calls which have waited to connect to an agent via ACD.

- *Number of calls answered (CA):* Total number of calls that are connected to ACD and eventually connected to an agent.

- *Average service time ($\overline{ST}$)*: Average time that an agent spends dealing with calls. This indicator may also be called as Average ACD Time ($\overline{ACD}$).

- *Average response time ($\overline{RT}$):* Average time that a caller spends waiting in the ACD queue before connecting to an agent.

- *Average waiting time before call loss ($\overline{WTBCL}$) :* Average amount of time that a customer spends waiting in the ACD queue before he/she abandons the call.

- *Average number of agents busy (AB):* Number of agents that are busy dealing with customers at time t.

- *Average number of agents logged off (ALF):* Number of agents logged off at time t.

- *Average IVR time:* Average amount of time that a customer spends before leaving the system or joining in the ACD queue.

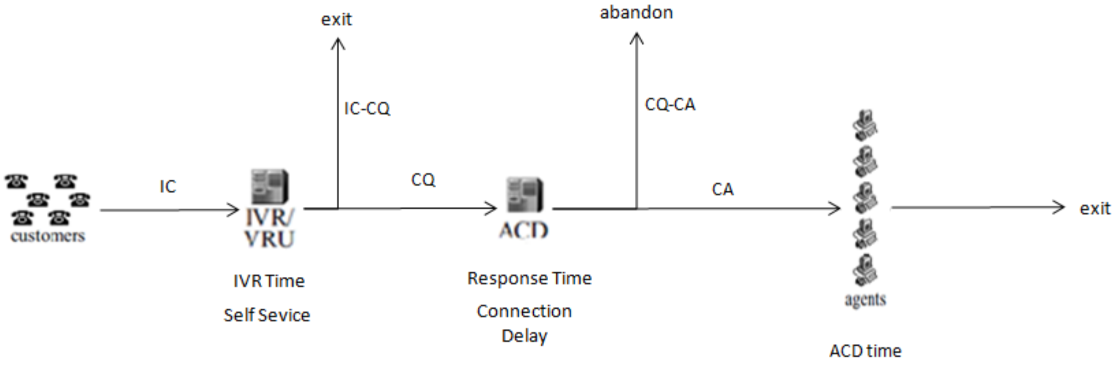

Figure 2.1. Call center performance indicators.

The second category which is related to individual performances includes agent statistics.

- *Number of calls received:* Number of calls that the agent dealt in a day.
- *Average time spent with a call:* Average handling time of a call for the agent.
- *Working Time:* Total amount of time that an agent is logged-on.
- *Amount of time the agent is logged off:* Total amount of time that the agent is logged off.
- *Amount of time the agent is busy with a call:* Total busy time for the agent.
- *Amount of time the agent is logged on and idle:* Total amount of time that the agent is ready to take a new call.

The six performance metrics will are included in our model will be presented in Chapter 3.

As it will be discussed in a more detailed way in the next sections, most of the studies on call center performance evaluations are on agent performances. The main reason for this is that call centers are highly labor intensive systems and agent based cost items compose about 60-80% of the overall operating budget (Ma *et al*., 2011). More examples of performance evaluation works will be given in the next parts.

2.2. Multivariate Statistical Process Control and Principal Component Analysis

Multivariate statistical process control (MSPC) may be described as a collection of algorithms which are used to extract information from correlated data sets (AlGhazzawi and Lennox, 2009). Although these techniques may differ in terms of methodology, their main common point is to be based on multivariate statistics created by considering the underlying correlation structure. While the original variables are used in computational steps of some of them, others have a methodology constructed on identified artificial variables replacing the original ones. The most common multivariate MSPC techniques using original variables are, Hotelling's T^2 (Hotelling, 1947), Multivariate Cumulative Sum (MCUSUM) proposed by Crosier (1988) and Multivariate Exponentially Weighted Moving Average (MEWMA) proposed by (Lowry *et al*., 1992). As stated by (J. Surtihadi *et al*., 2004) all these multivariate control charts which are extensions of their univariate forms have their weaknesses about uniquely detecting the shifts in the covariance matrix. Another major drawback of these multivariate control methods is their inability in detecting the reasons of produced signals. Although some methods concerning this point were already introduced such as Mason Young Tracy (MYT) decomposition proposed by (Mason *et al*., 1995) which is mainly based on decomposing the charting statistic, or an alternative approach proposed by (Alfaro *et al*., 2009) treating the issue as a classification problem, they all had their common drawbacks in practicality and efficiency.

Furthermore, another potential disadvantage of these control methods is that they are constructed on multivariate normality assumption which is usually difficult to validate and mostly inappropriate (Chang, 2007). Additionally, a basic assumption of these traditional multivariate process control methods is the independency of variables over time. However, most real life processes have an inherent dynamic nature producing various types of autocorrelated observations. It was observed in researches such as (Mastrangelo *et al*., 1996), (Lu and Reynolds, 1999) and (Lu and Reynolds, 2001) that when the autocorrelation values are low and process shifts are not large, these traditional multivariate methods demonstrate fairly acceptable performances with autocorrelated data sets. However, it is evident that an increased robustness is required for the analysis of most real life applications which usually have highly autocorrelated structures.

The usual practice so as to deal with autocorrelated nature of the data is to use the residuals resulted by a time series model fitting the data set. When a correct model is fitted to the data set, the residuals are consequently uncorrelated over time and they turn out to be the variables which will be used in constructing the process control chart. However, this approach has its drawbacks mostly in inferential steps. One of them is about the stationarity concept. Although, for the cases where the analyzed process is a stationary one, it is easy to conclude for a signaling point that it is the mean or the variance of the process that significantly changed, this inference is not so straightforward in cases where the process is nonstationary. Since a nonstationary process has no constant mean or variance, it is a hard task to understand whether this is a real abnormal condition or a step change caused by drifts in the parameters of the process. Moreover, (Runger, 2002) asserted that when the residuals are used, the performance charts may perform unnecessarily poor because of an incorrect reduction in the magnitude of a signal. This argument has clearly appeared in pre-application experimental steps of our study. Since the T^2 values created by using the residuals resulted by related ARIMA model, have highly decreased compared to significance limits, the number of out-of-control points incorrectly reduced.

Another approach for dealing with autocorrelation is the adjustment of control limits and re-estimation of process parameters for autocorrelated data. However this approach is not expected to perform well under high levels of autocorrelation.

As stated in the first paragraph of this section, an alternative method for multivariate process monitoring is to replace original variables with artificial ones which are called latent variables. Two major techniques using latent variables are Principal Component Analysis (PCA) and Partial Least Squares (PLS) methods. Despite the fact that they both rely on singular value decomposition of a given data set, these two methods have fairly different algorithms. While the PLS method is a multivariate regression tool, PCA is an unsupervised data analysis technique (AlGhazzawi and Lennox, 2009). When the data set involves both input and output variables, PLS is an appropriate way to construct the latent structure. If only input or only output variables are included, PCA is a more preferable way as an easier to implement method (Mastrangelo *et al*., 1996). Performance metrics that will be used in our study are introduced in Chapter 3. Since all the variables that will be included in our model are output variables, the proposed method is based on PCA. A brief description of PCA method will be given in incoming parts. For more details on these methods the reader is referred to (Jackson, 1991) for PCA and to (Geladi and Kowalski, 1986) for PLS method.

Principal Component Analysis is a multivariate technique in which a set of correlated variables are transformed to another (hopefully) smaller set of variables which are orthogonal to each other (Jackson, 1991). The objective of PCA is a dimensional change by transforming the original variables into a new set independent variables which are actually some linear combinations of the original ones. These new independent variables are called principal components. A nonsingular and invertible covariance matrix is the initialization point of the PCA method (Jackson, 1991). Assume that we have a *n x p* data matrix X where n is the number of observations, p is the number of variables and S is the covariance matrix of these p variables. As it is valid for all multivariate control procedures, if the variables are in different units or if there is big variance differences among them, the standardized form of variables are used in computations. Obviously, the covariance matrix of the standardized variables is the same as the correlation matrix of the originals ones. Consequently the covariance matrix may be replaced by the correlation matrix in some cases. In this study, S will be called as covariance matrix and X will represent the standardized form of variables. Having specified this detail, the principal components correspond to the eigenvectors of S. The PCA decomposition of X is as follows:

$$X' = U_k * z + E \tag{2.1}$$

in which, k is the number of principal components retained in the model, U is the *p x p* matrix containing the principal components as its columns, U_k is the *p x k* matrix containing the first k columns of U, z is the *k x n* matrix containing the principal component scores and E is the *p x n* matrix containing the standardized residuals.

There are number of ways to decide about the number of principal components to be retained in the model. Methods such as SCREE test, percent variation explained and cross-validation are the most commonly used ones. In our study, the percent variation explained method will be used. This method is based on the percent of accumulated variances of principal components which are determined by eigenvalues of the matrix S. Reasons for this preference will be given in Chapter 4.

For almost all examples of PCA applications, k is smaller than p meaning that a smaller number of underlying characteristics actually drive the process (AlGhazzawi and Lennox, 2009). Obviously for values of k smaller than p, the error terms in Equation 2.1 will be different than zero since the computed X matrix does not contain the exact values but the predictions of the original ones.

Two types of charting statistics which are T^2 and Squared Prediction Error (SPE) are traced simultaneously within PCA method. The T^2 statistic is about the distance of an individual observation from an established standard. As (Jackson, 1991) stated the important point about an outlier in T^2 chart is that it would be still detected as an outlier if conventional multivariate techniques had been used instead of PCA. However, by the use of PCA the probability of detecting it and possibilities to sort out the reasons of this extremity is enhanced. T^2 value of a particular observations is computed as in Equation 2.2.

$$T^2 = z' * L^{-1} * z \tag{2.2}$$

where z is the k x 1 vector containing the principal component scores of this particular observation, and L is a diagonal matrix containing the first k eigenvalues of S as its diagonal elements.

Since T^2 is a squared value it has only an upper bound. The distribution of T^2 is related to the F distribution by Equation 2.3.

$$T^2_{k,n,\alpha} = \frac{k(n-1)}{n-k} * F_{k,n-k,\alpha} \tag{2.3}$$

where α is the significance level.

The second charting statistic is SPE. As the result of dimension reduction, the PCA method is unable to detect shifts in the process mean vector that are orthogonal to the space spanned by the first k principal components. To overcome this defect it is essential to track *SPE* values besides T^2 values. Two results may be extracted from a signaling point in *SPE* chart; A shift in the mean vector orthogonal to the spanned space have happened or an unexpected change in the correlation matrix have occurred.

SPE value for a particular observation is simply the sum of squares of standardized residuals of related variables. That is, sum of squares of i^{th} column variables of E matrix is the *SPE* value for i^{th} observation. Let e_i denotes the i^{th} column of E matrix, then *SPE* value for i^{th} observation is

$$SPE = \|e_i\|_2^2 \tag{2.4}$$

Similar to T^2, *SPE* values have only an upper bound, however since it is not directly related to a known distribution; its upper bound computation is fairly different. The upper limit for *SPE* is

$$SPE_\alpha = \theta_1 \left[\frac{C_\alpha \sqrt{2\theta_2 h_0^2}}{\theta_1} + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} + 1 \right]^{1/h_0} \tag{2.5}$$

where C_α is the normal deviate cutting of an area of α under the upper tail of standard normal distribution if h_0 is positive and under the lower tail of a standard normal distribution if h_0 is negative. θ_1, θ_2, θ_3 values are the sums of first, second, and third powers of the variances of unused principal components, respectively. And the h_0 value is computed as follows

$$h_0 = 1 - \frac{2\theta_1\theta_3}{3\theta_2^2} \tag{2.6}$$

For further details about Equation 2.5 the reader is referred to (Jackson, 1991).

In traditional multivariate process monitoring with PCA, two phases of operations are conducted. The first phase is about determining the historical data set (HDS) representing normal operating conditions of a process. While this phase may be performed applying a PCA to the selected process period and eliminating the out-of control points in T^2 and *SPE* charts, there may be some cases that it should be performed based on pure process experience. Once the HDS is determined, the phase two operations begin. This is the phase in which the rest of the process is evaluated depending of the normal operating conditions set by the HDS. Deviations from these conditions are reflects to T^2, *SPE* charts as signaling points. Obviously, determining the reasons of these upset conditions is essential for a monitoring system. At this point, it is required to examine the contributions of each variable for a signaling point. Examination of these contributions is more straightforward for *SPE* charts than it is for T^2 charts. As stated earlier, the *SPE* values are simply the sum of squares of standardized residuals of related variables. Therefore, it is sufficient to check the standardized residuals in order to understand which variable or variables cause *SPE* value to significantly increase. On the other hand, because the T^2 values are not computed directly from original variables but from principal component scores, it is a harder task to reveal the causal link between the signals and variables. In our study, the *D* statistic inspired by the U^2 statistic proposed by (Runger *et al*., 1996) is used to handle this decomposition problem. Details about this statistic will be given in next chapters.

One of the reasons why PCA is preferred for this study to other conventional multivariate process monitoring methods is that the assumption of multivariate normality is not required when PCA is applied. As stated earlier, this assumption is valid for all conventional multivariate methods and mostly not satisfied. Moreover, autocorrelated structure of a process commonly means the violation of the main underlying assumption of traditional methods which is independency of observations. (Mastrangelo et al., 1996) revealed in their study that traditional multivariate control chats based on T^2 statistic are expected to be less sensitive to autocorrelation compared to univariate techniques. However, in the same study they enclosed the argument that using latent variables for autocorrelated processes performs remarkably better in indicating process upsets. Furthermore, through clearer structure of its T^2 and *SPE* statistics, PCA method provides a flexibility to the user at inferential steps. With the help of easy to understand contribution plots, revealing of causal links between signals and variables becomes more straightforward and this makes it possible to explicitly detect correlation related performance upsets.

On the other hand, the major drawback of PCA is its time-invariant nature. This deficiency makes the method prone to cause some serious problems in analyzing nonstationary processes. The stability of HDS translates into a monitoring method depending on a stable mean, variance and correlation structure to assess the rest of the process. However, this case may mislead decision makers in cases where the process parameters are actually changing in time as a characteristic of the process. With the purpose of covering this point, various kind of adaptive PCA models were introduced in the literature. (Wold, 1994) constructed an exponentially weighted technique using the complete data set to renew the model. (Li *et al*., 2000) proposed two recursive algorithms for adaptive process monitoring with PCA. Apart from adaptive capabilities of the proposed methods, the computational efficiency of the algorithms was of interest for their research. (Lane *et al*., 2003) presented the application of an extended version of (Li et al., 2000)' s study in a polymer film manufacturing process. Recently, (AlGhazzawi and Lennox, 2009) introduced kind of a different adaptive approach renewing the model using a stable-size moving window. They also used the same algorithm to create an adaptive PLS monitoring system.

Details of our proposed "Periodic Weighted Moving Average PCA" algorithm will be presented in the Chapter 4.

2.3. Performance Evaluation Applications

In this study, we propose an adapted method which has been used for industrial systems for performance analysis of call centers. As (Esposito and Tenenhaus, 2008) stated, the main description for performance analysis is a systematic management approach to find out what to do in short reaction times and collecting formal and informal data as a tool to set and meet organizational goals. At this point, performance evaluation is not a matter of simple data analysis but requires advanced statistical tools. Section 2.3.1 focuses on recent uses of statistical process control techniques in performance evaluation with various application fields and in Section 2.3.2, more specifically; call center performance evaluation studies are examined.

2.3.1. SPC Techniques in Performance Evaluation

In the early years of statistical process control techniques in performance evaluation, the studies were more focused on industrial examples (Funch and Kenettt, 1998). However, as the benefits of such methods becoming obvious, the application fields have enlarged as well and sectors such as environment, education, finance, and government performance evaluation projects have joined in application fields of SPC techniques.

(Corbett and Pan, 2002) used SPC techniques in evaluating environmental performance of an industrial system. In their paper, they build a detailed quantitative procedure using univariate CUSUM charts with process capability indices to evaluate environmental performance in terms of the risk of non-compliance situations arising. While modified CUSUM charts are monitoring the amount of emissions produced, with the collaboration of process capability analysis they create a tool to assess the likelihood that the process keeps meeting the legislative requirements.

In addition to various process performance evaluations, individual performance assessment has also covered by the application area of statistical process control. In 2006, within The National Assessment of Educational Progress (NAEP) which is an ongoing survey on the performance of school students in U.S., (Davier *et al.*, 2006) presented their proposed methodology which depends on a complex latent regression model and sampling weights about individual performance assessments of a hundred thousand examinees. As another example for the application of statistical approaches to individual performance measurement, (Shehab *et al.*, 1997) used several statistical process control charts in addition to a currently used metric to evaluate employee performances based on readiness to perform tests (RTP). As the result of their work, they proved that SPC techniques result in a more efficient and accurate performance compared to the currently used evaluation method. (Purcell, 1994) used SPC charts as a device to visualize and analyze individual performance in decision-making tasks. By submitting the data collected about the effects of order of training and experience on working, he managed to come up with a tool to make robust measurements for performance indicators of an individual in a computer supported environment.

(Leu and Lin, 2008) introduced SPC techniques into earned value management (EVM). Their main objective was to make up the shortage of traditional EVM technique in determining the performance variations of a project and related causes. Analysis of an ongoing project was made possible by the use of control charts allowing to determine possibly required preventive actions in a timely manner. Also, a concrete tool was allocated for the use of decision makers to assess possible project performance behaviors at the early stage. Moreover, (Qing *et al*., 2006) proposed simultaneous use of earned value management and statistical process control techniques in cost and schedule monitoring which they qualified as “The most important supportive activities” for the success of a process.

Another innovative application of SPC techniques was proposed in business activity monitoring area by (Au *et al*., 2007). With the developed model, an SPC based framework for modeling and tracking customer behaviors was intended. The proposed model involves several profiling algorithms to assess different behavioral patterns such as business changes, structural breakdowns and unwanted errors. The main statistic that their model depends on is exponentially weighted moving average (EWMA). The enhanced model is created by comparing the built algorithms in a simulation and in a customer churn detection example. The model facilitates executives to construct more robust and more efficient customer loyalty programs for churn reduction and fraud detection.

Moreover, financial performance measurement has also been in the application field of SPC techniques. As the result of a test application of SPC on a cross border remittance process in a Taiwanese bank, it was proven that an efficient adaptation of SPC in financial institutions is possible. (Tsai and Ou-Yang, 2010) intended to follow the success of this pilot application by working on three more work processes of the bank. The main goal of the study was to improve the compatibility of SPC in financial processes. At the end of the study, significant reductions in the averages and standard deviations of cycle times were stated as the benefits of the roadmap constructed depending on the applied model. Furthermore, practicability of getting the solutions via the examination of process capability and process stability indices was explicitly stated as a factor increasing the efficiency of the constructed model.

Additionally, (Rex, 1999) suggested the use of SPC techniques as a management tool in healthcare systems. SPC methods were used to monitor the process of outpatient service delivery. Depending on weekly ratings of functioning provided by the client and therapist, global outcomes for clients were computed. The linear change in functioning derived from these data was monitored in control charts using moving average method. After the determination of variability causes and revision of control limits, uncontrollable variation was removed from the service process. As the result of this research, it was indicated that the intended levels were not actually achieved in terms of some discharges and the mean level of client outcome. Thus, it was proven that some revisions were actually needed for the on-going process.

More particularly, the PCA method also has a large application field involving various disciplines. (Miletic *et al*., 2004) made online applications of multivariate statistics on manufacturing processes. Their applications including fault detection models for various manufacturing processes and a predictive modeling for a desulfurization process were basically depending on PCA and PLS methods. With a similar point of view, (AlGhazzawi and Lennox, 2009) developed two types of model predictive control monitoring tools employing PCA and PLS for a chemical process. Their model is used in real time to identify abnormalities in the process. In their work, they developed a new approach to static PCA and PLS models creating “Recursive PCA and Recursive PLS” to deal with the nonstationary nature of the data. As the result of both works very successful and easy to understand performance monitoring methods were constructed.

Another application of PCA method for performance monitoring in chemical industry was made by (Liang *et al*., 2006). In their study, two different chemical processes which are a two input by two output distiller and a heavy oil fractionator were examined. The proposed method was applied on both systems and resulted to be a monitoring method with increased sensitivity and perspicuity compared to conventionally used techniques.

Another innovative example of PCA applications is in aviation discipline. (Mujica *et al.*, 2008) presented an application of multi-way PCA in structural health monitoring of a commercial aircraft wing flap. In their work, they used multivariate methods on a damage identification problem considering the reciprocal relationship between sensors monitoring the structures.

Furthermore, (Li *et al.*, 2008) proved that multivariate process monitoring methods may turn into a valuable decision support tool for price makers in energy markets. By using PCA and cluster analysis in a collaborative way, they created a model to solve customer classification problem of a power supply network. Finally, as another example from energy industry (Ming *et al.*, 2009) combined PCA and backpropagation (BP) neural networks within a risk evaluation project in power supply industry. Through their risk evaluation model profiting from learning ability of neural networks and dimension reduction ability of PCA, they came up with a decision support system including historical experiences and avoiding subjectivity.

Having sum up various applications of statistical process control methods, call center performance evaluation works will be introduced in the next subsection.

2.3.2. Performance Evaluation Works for Call Centers

Performance of a call center is usually evaluated in two parts which are quality of service and efficiency. While the efficiency concept is often referred to agent utilization, quality of service may be differentiated as probability of blocking a customer, and customer satisfaction. Blocking a customer means connection delays. Customer satisfaction concept includes two main topics which are first call resolution and customer-agent communication. The first call resolution is a relatively newly born concept meaning the customer has a satisfying response at his/her first call and does not need to re-call for the same problem. As seen in performance type classifications, both quantitative and qualitative approaches are possible and in fact necessary for a proper performance evaluation in call centers.

In order to construct a healthy performance evaluation process, determining right measures and related objectives is a vital point. (Cleveland, 2007) stated seven measures that should be focused on by any kind of call center. These measures are strategic value, customer satisfaction and loyalty, employee satisfaction, quality and first-call resolution, service level and response time, forecast accuracy, and schedule adherence. He also implied that high-volume data produced by call centers is the initial significant problem that an executive has to confront. Data mining procedures are the common methods to deal with complex data sets. (Marcin *et al*., 2004) introduced applications of several data mining methods such as linear neural networks and probabilistic neural networks in assessing the quality of service in call centers. Their work had two main objectives. First one was a comparison between different models built on different techniques and the second objective was a sensitivity analysis to uncover relationship between the performance evaluation process and the actual performance. By this way, the authors intended to prove that performance evaluation is not only about assessment but it also has a big potential to increase individual and process performances. Another example of data mining applications in call centers was introduced by (He *et al*., 2007). In their work, data mining methods combined with statistical process control techniques was used to build a model for continuous quality improvement concept. Separate performance examination for IVR process and individuals were a key point for their research. Moreover, (Jing and Min, 2010) used a neural network algorithm to compute service quality grade of a call center depending on managerial requirements of the enterprise. At the end of the research, they proved that data mining offers a scientific and rational basis for managerial decision making steps.

Customer satisfaction and retention is another aspect of a call center’s performance. Based on recent studies, customers are becoming more and more demanding against call centers (Jaiswal, 2008). This case makes it harder for call centers to make their customers satisfied with the service provided. (Elaine, 2010) asserted that the biggest opportunity for call centers to retain their customers is in the first year of interacting with a new customer. He went on to add that service delivery has a close relationship with customer satisfaction which is one way to positively impact customer retention. At this point, the linkage between customer satisfaction and service quality is worth to analyze. (Dabholkar *et al*., 2000) asserted that although they are two closely related issues, an explicit examination of service quality and customer satisfaction is very useful for understanding customers' perception of provided service quality. As a more integrated application for customer satisfaction improvement, (Serra and Martucci, 2003) discussed the application of customer relationship management (CRM) for call centers. In their study, they divided CRM into three parts as operational, analytical and collaborator. They included that a properly applied CRM is a crucial point in obtaining customer satisfaction and loyalty.

Moreover, employee retention which is in turn obtained by employee satisfaction is another influential concept on performance of a call center. Recent studies have shown that an increase in employee satisfaction directly translates into an increase in customer satisfaction. (Ward, 2006) developed a mathematical model for analyzing the contribution of increased employee retention to call center performance. His work is based on an idealized model of a contact center with a stable agent number meaning that an immediate replacement occurs after an agent departure. In his proposed model, employee's performance is considered as a function of agent experience and the overall performance of the call center. On the other hand, the performance of the call center is deployed on agent retention distribution and employee performance function. Both of these concepts are assumed to be affected by management decisions. Examining the results given by constructed model, it is possible to describe the consequences of managerial actions on long-run average staff experience and long-run average call center performance. Additionally, as another study focusing on employee side of call center process, (Ramseook-Munhurrun *et al.*, 2009) discussed the service quality of a call center as perceived by the employees. The main purpose of their work was to reveal the factors determining agent satisfaction and behavioral intentions referring to willingness to recommend the call center and to stay. The proposed model that uses SERVQUAL method was constructed to measure both perception and expectation levels of agents. Additionally, it uses regression models to check the influence of service quality on employee satisfaction and behavioral intentions.

Furthermore, another and relatively newly arising concept in call center performance evaluation is the first call resolution. This concept has recently been hardly recommended by many researchers to be paid more attention. They argued that quality is more important than speed which is an important but not the leading indicator of excellence in service. The first call resolution is directly related to individual agent performance. However, despite this close relationship executives should very carefully evaluate this criterion since some factors affecting first call resolution rates are not controllable for agents. Supporting this argument, (Cleveland, 2007) said that "Components that lead to first-call resolution should also be built into specific quality objectives for agents (however, because not all aspects are within their control, these components must be selected carefully)." Various ways such as quality monitoring samples, customer information systems, call coding, direct questions asked to customers etc. may be used to track first call resolution rates.

Additionally, in terms of stochastic approaches to performance evaluation in call centers, simulation and queuing models are the most commonly used techniques. (Mehrotra and Fama, 2003) introduced the challenges and opportunities of simulation modeling for call centers. They state the multiple caller channels, increased complexity in call traffic and operations caused by acquisition/merger activities and outsourcing options as the main reasons making simulation modeling attractive. As the result of their work, they have put forward the flexibility and consideration of system dynamics as the most significant benefits of simulation modeling. Moreover, (Zhu and Zhu, 2009) considered a call center call handling process as a M/M/s/k+G queue model to assess process performance. At the end of the study, they proposed five concepts as performance criteria that decision makers may use in managerial decisions. These concepts are the average number of customers in waiting and service area, average number of busy servers, average number of customers in retrial area, probability of blocking and probability of loss.

Also, (Huo and Dai, 2008) added the call abandonment phenomenon into their model considering the call handling process as a markov chain. They implemented their model on a multi-skill call center process. Their work includes also a comparison between simulation modeling and proposed queuing model in terms of results and efficiency. As the result of the study, they showed that the proposed queuing model is able to obtain very similar results to that of simulation method in a significantly shorter time. Depending on comparisons made between with/without abandonment models, they additionally asserted that the abandonment has a quite large affect that it should be necessarily included in the queuing model.

As a totally different perspective, the dependability level of performance measurement results is another research area in call center performance evaluation. The reliability of assessment results is of main importance for decision makers. (Kornyshev and Duz, 1990) focused on accuracy of evaluating service quality in call centers. The main argument put forward in the article is “The accuracy of the measurement results depends on the duration of the measurement period”. Their proposed method to optimally decide on the compromise between the accuracy and duration of measurements depends on influences of measured random variables on network design, operation and maintenance.

To summarize, performance evaluation for call centers has both quantitative and qualitative aspects. Two key concepts are at issue for a successful performance evaluation; first one is to have properly defined performance indicators and the second one is the usage of a reliable performance measurement tool. Also, it was concluded that another important point for call center performance measurement is to be able to properly differentiate between individual performance and general process performances. Since, in some cases, the combination of these two may mask some problems and mislead the decision makers.

3. METHODOLOGY

This chapter will introduce the methodology followed in developing the performance evaluation model for an inbound call center process. The model development procedure involves the following four basic steps.

(i) Description of the call handling system
(ii) Examination of current performance evaluation technique and determination of its shortages
(iii) Objectives and model development
(iv) Data collection and analysis
(v) Algorithm development and proof of concept

Details for these steps are given in the next sections.

3.1. Process Overview

This section includes the details of the on-going call handling process and the currently used performance evaluation technique. Application part of our proposed method will be performed in a bank's call center. The call center is composed of two departments; inbound department and telemarketing department. In telemarketing department, calls are created by call center agents in order to commercialize banking products related to specific customer needs. That is, outbound calls happen within this part of the call center. The proposed method in this study will be applied to the inbound department. In the inbound department, calls are initiated by customers and following the steps displayed in Figure 2.2, an interaction between customers and agents begins. As stated earlier, there is no skill-based grouping among inbound agents, so all calls join in the same queue and distributed to available agents with the first-in-first-out criteria except the reports for stolen items and credit card passwords. Agents are grouped as teams under the supervision of a team leader. Depending on the workload fluctuations during different sections of a day, some of the telemarketing crew is subject to join to inbound agents. This action is called as "In-day resource allocation". Decisions about "In-day resource allocation" are manually handled based on the experiences of decision-makers.

During the conversation between agent and customer, the agent has individual records of the customer via CTI system on his computer screen. Within these records, if the sales pop-up system is on, the agent has different sales options that this customer may particularly be interested in. Whilst this conversation; a big percent of the agents which are called "sale groups", are responsible of proposing these sales to the customer. If the agent is not included in a sale group, no sale proposals are made. In case of increasing queue lengths and response times, this sales pop-up system is subject to be turned off. Decisions about turning off the system are made by team leaders working in a collaborative way with each other. This decision process has not an automated or an algorithmic structure but it is manually handled based on experiences and forecasting abilities of team leaders. Also, in some cases the customer may be put to on-hold status by the agent. While a customer is at on-hold status no conversation happens between agent and him/her until the agent gets back in touch after he/she finishes his/her work. These cases which are a variation reason for ACD times usually happen when an agent has to work out some small things on the computer or when he/she has to consult to an official about any issue. Below is the presented interaction between a customer and a sale-group agent.

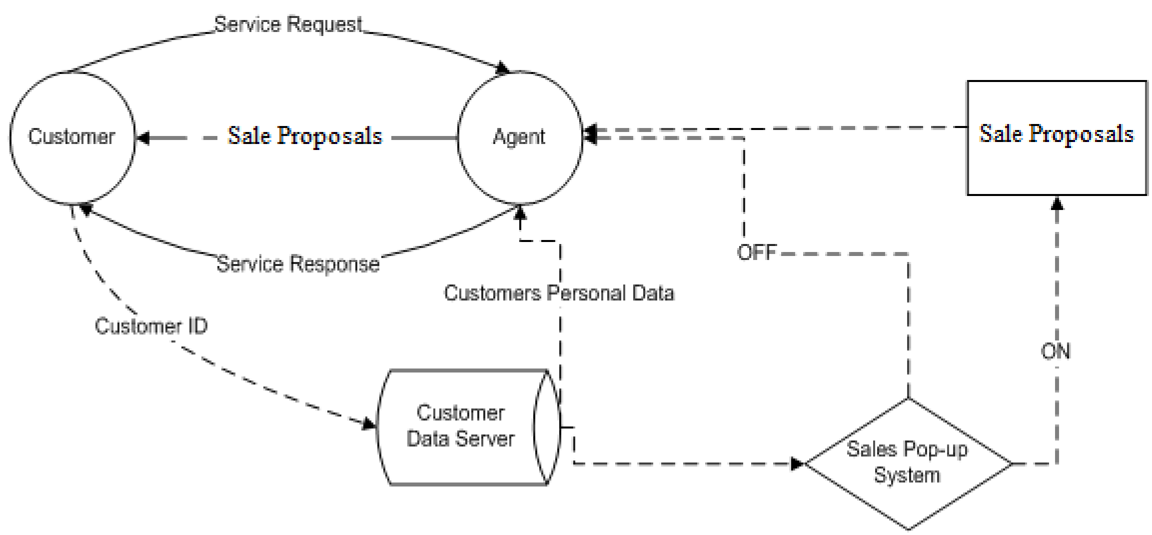

Figure 3.1. Interaction between a customer and a sale-group agent.

During a day, agents naturally have some of their working times logged off. Under normal circumstances, objective for logged off times is 7 minutes per hour in average. These logged off times are usually about relatively long lunch breaks, some small resting periods etc. Also, some of the agents might have some medical problems which force them to work with longer and more frequent breaks.

There are two shifts in a 24-hour day as day shift and night shift. In spite of the fact that the expected daily working time for an agent is about 8 hours, this value may greatly vary from one agent to another due to various reasons. The main reason for this variation is that part-time agents account for a big percent of call center agent crew. Also, in-day resource allocation is another reason for this fluctuation in working times. Due to such factors, there may be cases where some agents working 3-4 hours or even 1 hour in a day.

In the context of the performance evaluation, the executives are mainly focused on individual performances rather than general process performance. Agents are evaluated within a systematic approach which considers an agent’s performance in two separate parts as operational efficiency and service quality. In this evaluation method, the weight of operational part is 60% while the service quality has its weight as 40%. Operational efficiency part lies on award and penalty points collected depending on meeting the objectives for predefined performance metrics. Some of these metrics are number of calls put on-hold, total amount of time left on-hold, number of unanswered calls, staff time and average talking time per call. In service quality side, the method depends on the transformation of qualitative attributes into quantitative ones. This transformation is mostly effectuated by listening to randomly chosen calls handled by the agent. Some of the indicators of service quality are customer complaints, customer appreciations and customer loss compensations.

For call center process performance evaluation, the main and the unique criteria for call center managers is service level (SL). Service level denotes the percent of queued calls which are answered in less than 30 seconds after they involve in ACD queue.

$$SL = Pr\,\{RT \leq 30\} \qquad (3.1)$$

Although there is not a strict goal for the SL, depending on the interviews made with call center officials, it was argued that values higher than 85% are considered to be “successful” and values between 80% and 85% are considered to be “acceptable”. Also, through the discussions made with executives, 93% was defined as the efficiency threshold for the SL. Values that are higher than this threshold will be considered as inefficiently high within our model.

In addition to tracking SL values, some graphics such as average ACD times, topic based dispersion of incoming calls, daily number of incoming and abandoned calls are followed and presented to high-level executives in monthly meetings, with the purpose of detecting observable causes of SL value fluctuations.

As we have substantially represented general working process of the call center that we will be working on, shortages of the current performance evaluation method will be given in the next section.

3.2. Problem Definition

Call centers which have formerly been considered as a cost center are currently perceived to be a profit center. This transition in perception brings together with it that efficiency should be the main point in performance assessments. As it was stated before, call center managers essentially focus on agent performance evaluations to assess their units. When it comes to look at the big picture, they find it satisfactory to track service level values. However in our point of view, general performance assessment for a call center deserves a more detailed and to the point approach. About the traditional approach, it is possible to talk about two weaknesses.

Firstly, with the conventional technique, the efficiency concept is mostly disregarded. The main point for a work to be called as efficient is that it is not on maximum but optimal values. A manager examining service level values of the last month will not obviously be discomforted seeing values like 97% or 99%. However, bearing in mind that if we have a goal of 90%, values that are significantly less than this goal as well as values that are significantly higher than this goal will decrease our efficiency. In such a case, within our approach, we find it useful to ask the question “Is it really necessary to catch values that high?”. Because, it is beyond contest that some different cost related gains may be created by holding a lower but still successful service level. These gains may be about more sales, less average working times or less number of agents. Consequently, by this new approach, apart from unfolding efficiency related problems, it will be also possible to make additional comments about the success of forecasting units which are outside the operational part of the call center process.

Secondly, in addition to detect performance state of a unit, another important objective for a solid performance evaluation is to designate the reasons for detected fluctuations. Detection period of reasons has an essential role in taking necessary actions to prevent performance variations. Within the conventional method, even if some univariate statistics are traced, it is not an adequate way to properly detect the reasons for these performance variations. Because of this inadequacy of the traditional method which was caused by disregarding the causal links between process statistics, it is probable to miss or improperly detect points requiring to be fixed in the process.

Furthermore, depending on examinations we have made on past call center statistics, we realized that the traditional approach is poor about determining some significantly good performances which are not directly visible. Again the main reason for that is the disregard of causal links between process statistics. As an example of this case, two different days with the same service level value of 87% but with significantly different sales numbers may be given. Obviously, for an assessment made by conventional technique, there will be no difference in perception between these two which will be both considered as successful ones. However, in our point of view which is based on efficiency, there should be a definite differentiation between them. Such examples considering different performance metrics may be expanded.

Additionally, with the current approach, it is unavoidable for call center executives to see the process performance results with a lapse in time they occurred. Obviously, an approach closer to real time will improve reaction times for decision makers.

To sum up, two weaknesses of currently employed call center performance evaluation method have been the point of origin for our study. Insufficiency about giving a satisfactory response to the question of “Is it possible to make this unit work more efficiently?”, and lacking capability of revealing clear reasons for significantly skewed performances are two main shortages of the current performance assessment process.

3.3. Objectives and Model Development

Two main drawbacks of the current performance evaluation method were stated in the previous section. In our point of view, the most important modification that should be made to overcome these shortages is the involvement of other properly defined criteria in the assessment process. In other words, we aim to transform this univariate assessment process into a multivariate one. Through this transformation, it will be possible to make more detailed and explanatory examination for performance drifts by analyzing the changes of these correlated criteria compared to each other. Also with the interpretation of relations among variables we intend to reveal circumstantial performance differences.

Furthermore, in the context of problems about detecting hidden performance drifts which were stated at the problem definition part, constructing a model which is capable of detecting not only poor but also prevailing performances is among the objectives of this study.

Additionally, depending on the results maintained through this new approach, a decision support table will be introduced. In this systematic table, most likely reasons of the detected performance drifts will be given.

In summary, the key objective of this work may be stated as to come up with a performance evaluation model which is capable of covering the drawbacks of the currently employed method.

3.3.1. Definition of New Performance Metrics

Definition of new performance metrics is a crucial step for the success of proposed performance evaluation method. As emphasized earlier, the main point of our proposed method is to transform current univariate performance perception into a multivariate one.

$$p_1(x_1) \Rightarrow p_2(x_1, x_2 \dots, x_n) \tag{3.2}$$

In Equation 3.2, it is seen the change in the performance function where p_1 is currently used one-variable function, p_2 is the new multivariate function and n is number of newly defined performance metrics. Several interviews on the main reasons of performance variations were conducted with call center executives. Depending on these interviews, main reasons for performance fluctuations were determined. These reasons may be summarized in seven items.

- Variable agent numbers
- Variable staff times
- Changing sale policies
- Exceptionally high workloads
- Variable call durations (ACD times)
- Variable log-out times
- Technical problems

In addition to determine variation factors, another important issue in performance metric definition step is the controllability of defined variables. New performance metrics should be defined in such a manner that they should be directly affected by decisions or work performances of the call center staff. This rule may obviously be slackened if one would like to include some variables representing customer behaviors in the model. The most practical solution to address controllability issue is to define new variables as ratios rather than amounts. By this way, it would be possible to eliminate not performance but forecasting related quantitative variations.

Furthermore, since significant cause and effect relationships would be beneficial within a multivariate approach in detecting the reasons of abnormal performances, this relational structure should also be considered for new variables.

Bearing these issues in mind, six new performance metrics that will be used in our model were determined. Although another user has the flexibility to change the time scale, performance assessments in this study will be in a daily basis. Consequently, all of our performance metrics will be given as daily values.

Below are six performance metrics that will be included in the proposed performance assessment model.

(i) *Answer Percentage (AP):* The percentage of customers in ACD queue who are successfully connected to an agent.

$$AP = \frac{CA}{CQ} \tag{3.3}$$

(ii) *Average ACD Time ($\overline{ACD}$):* The average length of the conversations between an agent and a customer. It may also be called "Average Service Time". This value will be stated in seconds.

$$\overline{ACD} = \frac{Total\ ACD\ Time}{CA} \tag{3.4}$$

(iii) *Average Response Time ($\overline{RT}$):* The average length of the period a customer spends waiting in the ACD queue to be connected to an agent. As a reminder, for callers who abandon the call this period is called Average Waiting Time Before Call Loss *($\overline{WTBCL}$)*. This value will be stated in seconds.

$$\overline{RT} = \frac{Total\ RT + Total\ WTBCL}{CQ} \tag{3.5}$$

(iv) *Call per Agent (CpA):* The ratio computed by dividing the total number of queued calls by the total number of inbound agents.

$$CpA = \frac{CQ}{Total\ number\ of\ inbound\ agents} \tag{3.6}$$

(v) *Sales per call (SpC):* The ratio computed by dividing the total number of sales made during inbound calls by the number of calls answered.

$$SpC = \frac{Total\ sales}{Total\ number\ of\ calls\ handled\ by\ sale{-}group\ agents} \tag{3.7}$$

(vi) *Average working time ($\overline{WT}$):* This is the average of logged-on time for inbound agents. This value will be stated in minutes.

$$\overline{WT} = \frac{Total\ working\ time\ of\ inbound\ agents}{Total\ number\ of\ inbound\ agents} \tag{3.8}$$

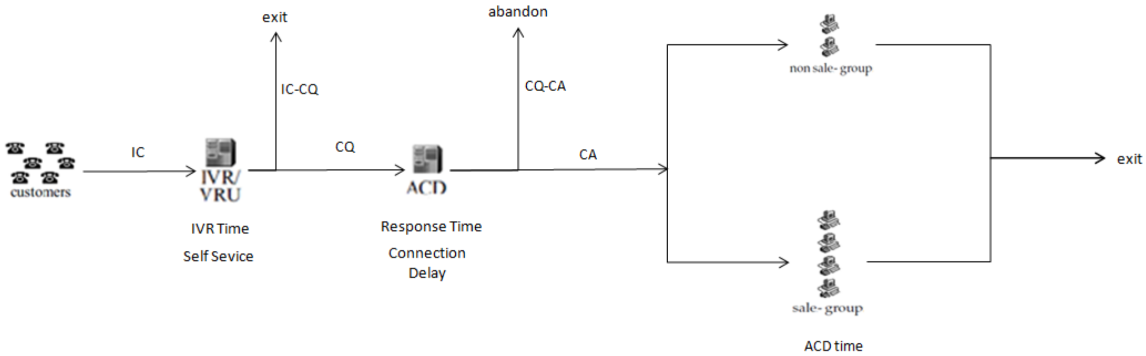

Figure 3.2. Current system illustration.

At the preliminary stage of determining these new performance metrics, “Average Occupancy” values representing the percent of staff time in which an agent is ready to handle a call, were also intended to be included with a view to embody the affect of variable log-out times in our model. However, it was observed that due to too small variances of this variable, number of unwanted signals increased. Since this would be considered as a false alarm by call center crew, this variable was excluded from the model to increase the likeliness that any alarm created by the method to be of interest to executives. On the other hand, inclusion of variables such as “Average After Call Works Time” and “Average Waiting Time Before Call Loss” would add value to inferential abilities of our model. However, because of technological constraints these variables could not be included.

Also, the $\overline{WT}$ value may be seen as an individual performance indicator and one may argue that it should not be included in a process performance evaluation model. However, we have already noted that the working times are flexible and the number of agents which is used to compute the *CpA* metric involves all the agents without considering the length of their staff time. For this reason, in order to prevent misleading results that may be caused by significantly drifted average working times, it was decided to add this statistic in the model.

Furthermore, as it is observed, the currently used unique performance indicator Service Level (SL) is intentionally not included in the model. The reason for this exclusion is that contrarily to other variables, it has a strict threshold which is 80% in practice. Consequently, within our proposed method which uses distributional properties of variables rather than deterministic values to designate extreme performances, this threshold would not be valid. That is, if a call center executive would like to have clear signals whenever the SL values go down 80%, the SL variable should not directly included in multivariate performance monitoring but should explicitly be traced. Obviously, service levels and average response times are closely related to each other. However, because of not symmetric but skewed distribution of response times and possible highly variable variances, it is not possible to talk about a strict linearity between these two metrics. It is important to note this point to be able to make clearer comments on possible incoherent changes of average response times and service levels.

3.4. Data Collection and Analysis

This part of our study is composed of three parts which are correlation, stationarity and distribution analysis of the collected data.

Two years of data were collected to be used in our study. First of all, the main reason that one requires to use multivariate process control methods is the presence of a significant correlation structure between variables. As the result of dynamic form of our process, the correlation structure of performance metrics is subject to change within time. However, it always holds its significance requiring to be analyzed within a multivariate approach. Below are correlation tables created by using arbitrarily chosen time sequences in two years.

Table 3.1. Correlations for months 1-4.

	% Answer	Average ACD. Time	Average Response Time	Call per agent	Sales per call	Average Working Time
% Answer	1	0.293	-0.8903	-0.1617	0.4387	-0.0248
Average ACD. Time	0.293	1	-0.187	-0.5115	0.6674	-0.1754
Average Response Time	-0.8903	-0.187	1	0.1764	-0.429	0.012
Call per agent	-0.1617	-0.5115	0.1764	1	-0.3219	0.7289
Sales per call	0.4387	0.6674	-0.429	-0.3219	1	0.0642
Average Working Time	-0.0248	-0.1754	0.012	0.7289	0.0642	1

Table 3.2. Correlations for months 5-16.

	% Answer	Average ACD. Time	Average Response Time	Call per agent	Sales per call	Average Working Time
% Answer	1	0.3913	-0.8887	-0.1563	0.4435	0.1046
Average ACD. Time	0.3913	1	-0.3636	-0.2709	0.7233	-0.218
Average Response Time	-0.8887	-0.3636	1	0.1774	-0.499	-0.0485
Call per agent	-0.1563	-0.2709	0.1774	1	-0.2012	0.4633
Sales per call	0.4435	0.7233	-0.499	-0.2012	1	0.0235
Average Working Time	0.1046	-0.218	-0.0485	0.4633	0.0235	1

Table 3.3. Correlations for months 17-24.

	% Answer	Average ACD. Time	Average Response Time	Call per agent	Sales per call	Average Working Time
% Answer	1	0.2947	-0.6082	-0.2033	0.2283	-0.0939
Average ACD. Time	0.2947	1	-0.1749	0.1954	0.6346	-0.138
Average Response Time	-0.6082	-0.1749	1	0.1902	-0.1019	0.0514
Call per agent	-0.2033	0.1954	0.1902	1	-0.1623	0.282
Sales per call	0.2283	0.6346	-0.1019	-0.1623	1	-0.085
Average Working Time	-0.0939	-0.138	0.0514	0.282	-0.085	1

As stated earlier, the starting point of PCA method is a non-singular covariance or correlation matrix. If two or more variables are perfectly correlated in a data set, this case is called “exact collinearity” and it may cause the covariance or correlation matrix to be singular (Mason and Young, 2002). In such a case, one of these perfectly correlated variables should be excluded from the data set. In order to check our data set against exact collinearities a guideline recommended by (Mason and Young, 2002) will be implemented. This guideline is based on condition indices which are described as the square root of the ratios of the largest eigenvalue of a correlation matrix to each of the other eigenvalues.

$$Condition\ index\ (i) = \sqrt{\frac{Largest\ eigen\ value}{i^{th}\ eigenvalue}} \qquad i = 1,2,\ldots p \tag{3.9}$$

where p is the number of variables.

The threshold for these indices is given as 30. A condition index which is greater than this value is considered to imply an exact collinearity in the data set. Table 3.4 includes eigenvalues and related condition indices for the correlation tables included in Table 3.1, 3.2 and 3.3.

Table 3.4. Eigenvalues and condition indices.

Months 1-4		Months 5-16		Months 17-24	
Eigenvalues	Condition indices	Eigenvalues	Condition indices	Eigenvalues	Condition indices
2.7256	1	2.7592	1	2.1075	1
1.6722	1.277	1.4965	1.358	1.3783	1.237
1.0504	1.611	0.8922	1.759	1.1392	1.360
0.2845	3.095	0.51	2.326	0.7674	1.657
0.1794	3.898	0.2491	3.328	0.3777	2.362
0.0878	5.572	0.093	5.447	0.2298	3.028

As demonstrated in Table 3.4, no exact collinearity is present in data set. Consequently the singularity of the correlation matrix is verified.

Secondly, as stated in the third chapter, although the conventional PCA has a proven effectiveness against autocorrelated structure of the data, because of its time-invariant nature, it has a major limitation in assessing nonstationary processes. At this point, autocorrelation functions of our two-year data will be examined. The autocorrelation functions for our six performance metrics may be seen in Appendix A.

As revealed by autocorrelation functions constructed by using the complete two-year data set, all of our variables clearly have nonstationary structures. This characteristic structure with consistently significant autocorrelation values is fairly sufficient to argue that the variables are nonstationary. Also, a more quantitative approach to decide about nonstationarity is to examine the ARIMA models fitting the data sets. The reader may refer to (Box and Jenkins, 1970) for details in time series analysis.

Moreover, considering the consistent spikes in the multiples of seventh lag it is easy to say that there is a weekly seasonality for all seven variables. With the additional examination of first differences of data, the seasonal structure becomes more apparent. The most common application to deal with seasonality is to rescale the original values with seasonal factors. However, with close inspection of the data, it was realized that the weekly seasonality is mostly caused by mean and standard deviation differences between weekdays and weekends. Moreover, it was also observed that particularly for weekend values the seasonal structure itself is subject to change within time. That is, while at some periods Saturdays have a tendency to have greater values than Sundays. This relationship may change vice versa in time. Taking all of these factors into account, it was decided to assess weekday and weekend performances separately. Obviously, as the result of this separation decision, deseasonalization operation will be applied separately to weekdays and weekends considering the period lengths of five and two, respectively.

At this point, it is required to re-examine the autocorrelation functions of variables for weekdays and weekends to see whether there are other seasonality structure changes among these two groups. Original and first difference autocorrelation functions of weekdays and weekends may be found in Appendix B. After this re-examination, it was observed that while some performance metrics show a nonseasonal form for both weekdays and weekends, others may have seasonal structures in at least one or both groups. Having detected this variant seasonality structures, the performance metrics, if necessary, are re-scaled depending on proper seasonal indices. In Table 3.5 and 3.6 seasonal indices for necessary variables are given with related monthly periods.

Table 3.5. Seasonal indices for weekdays.

	Call per agent	Sales per call	Average Working Time
Month numbers	1-24	1-24	1-24
Monday	1.043	0.911	1.016
Tuesday	1.004	1.021	1.007
Wednesday	1.002	1.006	1.004
Thursday	0.968	1.033	0.985
Friday	0.981	1.027	0.985

As seen in Table 3.5 and Table 3.6 in which month number rows represents related monthly periods, there are three variables displaying a seasonal change for both weekdays and weekends. While *SpC* values have seasonality in weekdays, no seasonality is observed for this variable among Saturdays and Sundays. Additionally, $\overline{ACD}$ values also show different characteristics for two day groups.

Table 3.6. Seasonal indices for weekends.

	Average ACD. Time		Call per agent		Average Working Time	
Month numbers	1-7	8-24	1-5	6-24	1-5	6-24
Saturday	0.991	1.021	1.119	0.935	1.019	0.933
Sunday	1.008	0.978	0.880	1.064	0.980	1.066

The original and first order difference autocorrelation functions of deseasonalized data demonstrating that the seasonality was discarded are attached in Appendix C. In the proposed performance evaluation method, the deseasonalized version of two-year data will be used. Future values which will be included in the model should similarly, considering the case of recent months, be rescaled depending on related seasonal indices in Table 3.5 or Table 3.6.

Finally, in the context of the distribution of the performance metrics, as stated in the literature review part, multivariate normality assumption is not essential so as the PCA method to perform well on indicating process upsets. Depending on this argument, the multivariate normality of the collected data will not be tried to be validated.

4. PROPOSED ALGORITHM AND ITS APPLICATION

This chapter focuses on the details of the proposed performance assessment algorithm and its application to the inbound department of a bank's call center. The algorithm basically depends on multivariate statistics. As it was mentioned in the earlier chapters, PCA has an essential role in our method for its high adaptability, robustness against autocorrelation and flexibility during interpretation steps.

As stated earlier, traditional MSPC applications depend on static principal component models. Such models employ a fixed reference data set representing a stable process nature. However, most of real life manufacturing or service processes have an inherent nonstationary structure. If one uses a static model to trace such processes, it is unavoidable that some process drifts which are in fact caused by the nature of the process will be observed as process abnormalities. In order to translate this case into our call center model, we should examine possible daily, weekly and monthly process drifts that result as the nature of the process.

In Chapter 3, the collected data were examined on daily and weekly basis and it was decided to evaluate weekday and weekend performances separately. Additionally, necessary seasonal adjustments were applied on both day groups. As the result of these actions, possible misleading results which may be caused by inherent daily and weekly process drifts were prevented.

In addition to daily and weekly drifts, interviews made with call center executives showed that six performance metrics are subject to significant changes from one month to another as the nature of call center dynamics. One of the reasons causing these drifts was stated as the large number of part-time student agents. *CpA* metric is the most probable one to be affected by this factor. Mid-term and final exam periods as well as summer vacations play a critical role in the availability of such student agents. Thus, a nonstationary model is naturally dictated. Also, it was included by the executives that possibly increased workloads have the potential to create step changes in performance metrics at the beginning and toward the end of a year. As such examples may be expanded, it turns out that using a fixed historical data set representing a stable mean, variance and covariance structure would not be logical and would in fact be misleading in analyzing such as system.

In order to cover this point, we propose a “Periodic Weighted Moving Average Principal Component Analysis” algorithm. The algorithm has a similar recursive logic with the ones proposed by (Li *et al*., 2000), (Lane *et al*., 2003) and (AlGhazzawi and Lennox, 2009). In the algorithms proposed by (Li *et al*., 2000) and (Lane *et al*., 2003), forgetting factors are used so as to adapt the model to process drifts. (Wang *et al*., 2003) and (Sandoz, 2003) proved that such algorithms employing forgetting factors are actually defective since they tend to easily adapt in abnormal conditions occurring in the process. On the other hand, in the algorithm constructed by (AlGhazzawi and Lennox, 2009), this forgetting factor approach is replaced by a stable-size window renewed at each time a new observation becomes available. It would not be appropriate to apply this approach into our system in a straightforward manner since although the process may have some monthly drifts, it would be misleading to assess any given day strictly depending on a particular number of recent days.

Our algorithm depends on two main parameters which are the “Dominant period” (DT) and “Weighting Factor” (WF). Dominant period is the number of most recent days of which contributions will be increased while recursively updating the mean, standard deviation and covariance structure. The weighting factor is a coefficient between 0 and 1. It is the weight of the relevant statistics of the dominant period. Consequently, the weight of the statistics for the remaining days becomes 1-WF. The parameter couple employed in a particular scenario will be represented as

$$\Phi(DP,WF).$$

The initial point of the methodology followed in the selection steps of two main parameters is that the nonstationary structure which should be traced by our model is monthly based. Consequently, the dominant periods are considered as multiples of the total number weekdays and weekends in a month which are 20 and 8, respectively. About the WF, it is evident that the DP should have a higher weight than the previous days.

Consequently, 0.6, 0.7 and 0.8 values will be included in the trial steps. Details about determination of these parameters will be given in the next sections.

In addition to main parameter estimations, selection of process period embodying the HDS is another problematic area for multivariate statistical process control. Construction of multivariate HDS is a more complicated task compared to the case of univariate procedures (Mason and Young, 2002). Although in recursive latent structures, the initially selected HDS is subject to be renewed as new observations become available, it still remains to be determinant for particularly preliminary group of observations. The nonstationary nature of the process is the main factor complicating the HDS selection step. For most of traditional stationary PCA applications, this step is a routine quantitative operation comprising the elimination of outliers from the selected data period by a static MSPC technique. However, for nonstationary cases, both quantitative and qualitative methods may be used to construct the initial HDS. In this study, depending on discussions with the call center executives, the HDS for weekday model was decided to be selected from within the first four months, 85 days, out of the available 24 months. Similarly, for the weekend model, the first six months, 48 days, were used to determine the HDS.

Visual inspection may be useful for the detection of outliers to create the HDS for relatively small data sets. However, it is not always effective. In this study, Moving Window Hampel Filter by (Pearson, 2002) was employed for the detection and removal of outliers. Hampel filter is an outlier-resistant substitute of outlier-sensitive mean and standard deviation estimates. This alternative filtering technique eliminates the outlier sensitivity by replacing the mean and standard deviation with median and median absolute deviation from the median (MAD). The MAD estimate S_{MAD} is defined as

$$S_{MAD} = 1.4826 * median\{|x_n - x^{\dagger}|\} \tag{4.1}$$

where x_n is a data sequence composing of n elements, and $x^{\dagger}$ is its median. The factor 1.4826 is a constant that makes the expected value of S_{MAD} equal to the standard deviation for normally distributed data. As (Pearson, 2002) stated, characterization of nonstationary data sequences is very much influenced by local anomalies however; a Hampel filter that considers the whole data sequence to asses a particular point, is concerned with global anomalies. In the same paper, it was stated that applying the Hampel filter in a form of a moving data window centered at the point in question, is a simple modification to address this problem. This filter has two tuning parameters: the moving window half-width parameter K (i.e. the Hampel filter will be applied to data sub-sequence x_{n-K} through x_{n+K} to calculate the clean estimate of x_n) and the rejection threshold t. Depending on these two parameters; if

$$\left|x_n - x^{\dagger}\right| > t * S_{MAD} \tag{4.2}$$

then the x_n is declared an outlier.

As (Pearson, 2002) stated, selection of these two tuning parameters is strictly application-dependent. In our case, large *K* values such as *K = 9* or *K = 10* performed fairly poor with *t = 2* or *t = 3*. With these parameter couples, despite some clearly visible outliers in data sequences, almost all days survived the filter. These results suggested us that smaller K values should be considered for better results. Applying the filter with *K = 2*, *t = 3* and *K = 3*, *t = 3* to weekdays, very similar results were gathered in both cases. Likewise, although not as close as in the weekday case, these parameter couples again performed similarly in the weekend model. Taking the benefit of additional visual inspection, the filter was decided to be applied with tuning parameters *K = 3* and *t = 3*, in both weekday and weekend models.

In this work, the moving window Hampel filter was separately applied to each of the six performance metrics. That is, for weekday model, 85 days for each performance metric was filtered and those that satisfy Equation 4.2 were declared outliers. Then, days of which at least one of the performance metrics is outlier were excluded from the model. The same methodology was followed to select the HDS for the weekend model.

As the result of this selection, 68-day and 39-day HDS's were constructed for weekdays and weekends, respectively.

In the proposed algorithm, apart from the update of mean, standard deviation and covariance structure, the number of principal components to be retained in the model is also renewed at every new observation. As stated in the second chapter, there are number of ways to decide about the number of principal components to be used. Cross-validation is one of the most commonly used ones of these techniques. However, because the data lose its representativeness for the process as it gets older, cross-validation is not a suitable technique for recursive PCA approaches (Li *et al*., 2000). For this reason, considering its propriety for recursive approaches, percent variance explained (PVE) method was preferred to be used in this study. It is also referred as explained proportion of trace. It is defined as

$$PVE = \frac{\sum_{i=1}^{k} l_i}{\sum_{i=1}^{p} l_i} \tag{4.3}$$

where l_i represents the i^{th} eigenvalue of the covariance matrix.

Usual limits for this value are 90 to 95%. (Lane *et al*., 2003) recommended determining this limit by applying the cross-validation technique to the HDS and using the resulting value for the rest of the process. However, this approach did not perform successfully for our case. The limit value obtained by applying cross-validation on the HDS was 98%. When we used this limit for the evaluation of future observations, the resulting number of principal components retained in the model has increased. This case resulted in a significant rise in false signals given by the *SPE* statistic. Thus, 90% limit which is found to be most appropriate will be used in the model. In next sections, details about how the two main parameters were selected will be introduced.

Finally, in order to make a more sensitive process monitoring possible, 0.99 and 0.95 significance limits will be used together in both T^2 and *SPE* charts. 0.99 level will serve as the control limit. On the other hand, 0.95 level will be called as warning limit. It should be noted that if a decision maker is interested in examining smaller drifts in performance, points which are over the warning limit would evidently be noteworthy.

4.1. Algorithm Description

As stated earlier, the proposed method has the conventional recursive logic for principal component analysis. It is based on the update of the mean, variance and correlation matrix. Additionally, the HDS is also subject to augment each time a day with in-limit performance values happens.

The update of the HDS relies on the addition of in-control days. An in-control day represents a day where there are no significant performance changes. So, each time a day that have both of its *SPE* and T^2 values within the control limits takes place; it is added into the HDS. Otherwise, it is excluded from the HDS and called as a signal. Obviously, in signaling days, there is a significant performance change, and they should be closely examined to uncover the reasons of this drifted performance. Since such extreme performances appear in forms of out of control *SPE* and T^2 values, the answers about the reasons of these fluctuations lie on close inspection of these increased statistics. In order to detect the contributing variables for a high *SPE* value, the algorithm simply checks the standardized residuals. For the analysis of out-of-control T^2 values, (Runger *et al*., 1996) proposed the U^2 statistic which is based on the difference between two T^2 statistics. Following a similar logic, we have created the *D* statistic to identify the contribution of each variable to an out-of-control T^2 value. The *D* statistics is identified as the difference between the T^2 value computed on the full data set, and the T^2 value computed on the data set excluding the variable of interest.

$$D_{i(j)} = T_i^2 - T_{i(j)}^2 \tag{4.4}$$

where T_i^2 is the T^2 value for the ith observation computed by including all six performance metrics, and $T_{i(j)}^2$ is the T^2 value computed by excluding the variable j. More specifically, $T_{i(j)}^2$ is computed in the same way as Equation 2.2 but using a *z* vector created by excluding the jth variable. As for the *SPE* contribution case, variables making the largest contribution to an out-of-control T^2 statistic are considered the most significant in determining the reasons of a given signal.

It should be particularly noted that in Equation 4.5 and 4.6, the number of days whose weights are increased is equal to DP+1. The reason for that lies in Equation 4.10. While computing the weighted covariance matrix in Equation 4.10, the current day is added to the last DP number of days of the HDS. So, the weighted covariance matrix on which the PCA is performed, is computed by increasing the weights of the last DP+1 days in the temporary data matrix. Therefore, to satisfy the parallelism between all updating equations, last DP+1 days are weighted in Equation 4.5 and 4.6 as well.

The steps of the algorithm proposed for the "Periodic Weighted Moving Average Principal Component Analysis" are as follows:

Initialization Steps:

(i) Define the historical data set (X_t). This can either be by using an appropriate data filter, or, for smaller data set, by visual inspection of most recent observations.

(ii) Calculate the weighted mean ($\bar{x}_t$) and standard deviation ($\bar{s}_t$) vectors; assigning the weight specified by WF to the mean and standard deviation values given by last number of days specified by DP and 1-WF to the values given by older ones.

$$\bar{x}_t(i,1) = (1 - WF) * mean\left(X_t(1{:}end - DP - 1, i)\right) + WF * mean\left(X_t(end - DP{:}end, i)\right) \quad i = 1,2,..,p \tag{4.5}$$

$$\bar{s}_t(i,1) = (1 - WF) * stddev\left(X_t(1{:}end - DP - 1, i)\right) + WF * stddev\left(X_t(end - DP{:}end, i)\right) \quad i = 1,2,..,p \tag{4.6}$$

where *end* is the index of the last element of the related row or column.

(iii) Standardize X_t using mean and standard deviation vectors created in step ii.

$$X_t = \left(X_t - 1_n * \bar{x}_t{}^T\right) ./ (1_n * \bar{s}_t) \tag{4.7}$$

where 1_n= $[1,1,\ldots,1]^T \in R^n$

Online day evaluation steps:

(iv) For a new online vector (x_{t+1}) comprising the performance metrics of the current day, standardize x_{t+1} using $\bar{x}_t$ and $\bar{s}_t$.

$$x_{t+1} = (x_{t+1} - \bar{x}_t) \,./\, \bar{s}_t \tag{4.8}$$

(v) Create the temporary data matrix Y_t as follows,

$$Y_t = \begin{bmatrix} X_t \\ x_{t+1}^T \end{bmatrix} \tag{4.9}$$

(vi) Calculate the weighted covariance matrix for Y_t; assigning the weight specified by WF to the covariance matrix resulted by the last number of days specified by DP and 1-WF to the one resulted by older ones.

$$S_{weighted} = (1 - WF) * S_{Y_t(1:end-DP-1,:)} + WF * S_{Y_t(end-DP:end,:)} \tag{4.10}$$

where : , when it is used without any indices at its sides, represents the whole elements of the related row or column.

(vii) Calculate the eigenvectors and eigenvalues for the covariance matrix defined in step vi.

(viii) Calculate the number of principal components to be retained in the model.

(ix) Calculate the principal component scores and T^2 statistic for the current day.

(x) Calculate the *D* statistic.

(xi) Calculate the standardized residuals and *SPE* statistic for the current day.

(xii) Calculate the warning and control limits for *SPE* and T^2 statistics.

If both *SPE* and T^2 statistics are within the control limits, continue to step xiii, otherwise go to step iv for the evaluation of the next day.

(xiii) Include the current day in the historical data set as in Equation 4.11.

$$X_t = Y_t = \begin{bmatrix} X_t \\ x_{t+1}^T \end{bmatrix} \tag{4.11}$$

(xiv) Calculate new weighted mean $(\bar{x}_t)$ and standard deviation $(\bar{s}_t)$ vectors in the same way as stated in step ii but depending on updated X_t. Then, standardize X_t using these new $\bar{x}_t$ and $\bar{s}_t$ values.

(xv) Go to step 4 for the evaluation of the next day.

MATLAB commands of this proposed algorithm may be found in APPENDIX D.

Having given the details of our recursive PCA algorithm, the proposed multivariate process performance evaluation method may be summarized as in Figure 4.1.

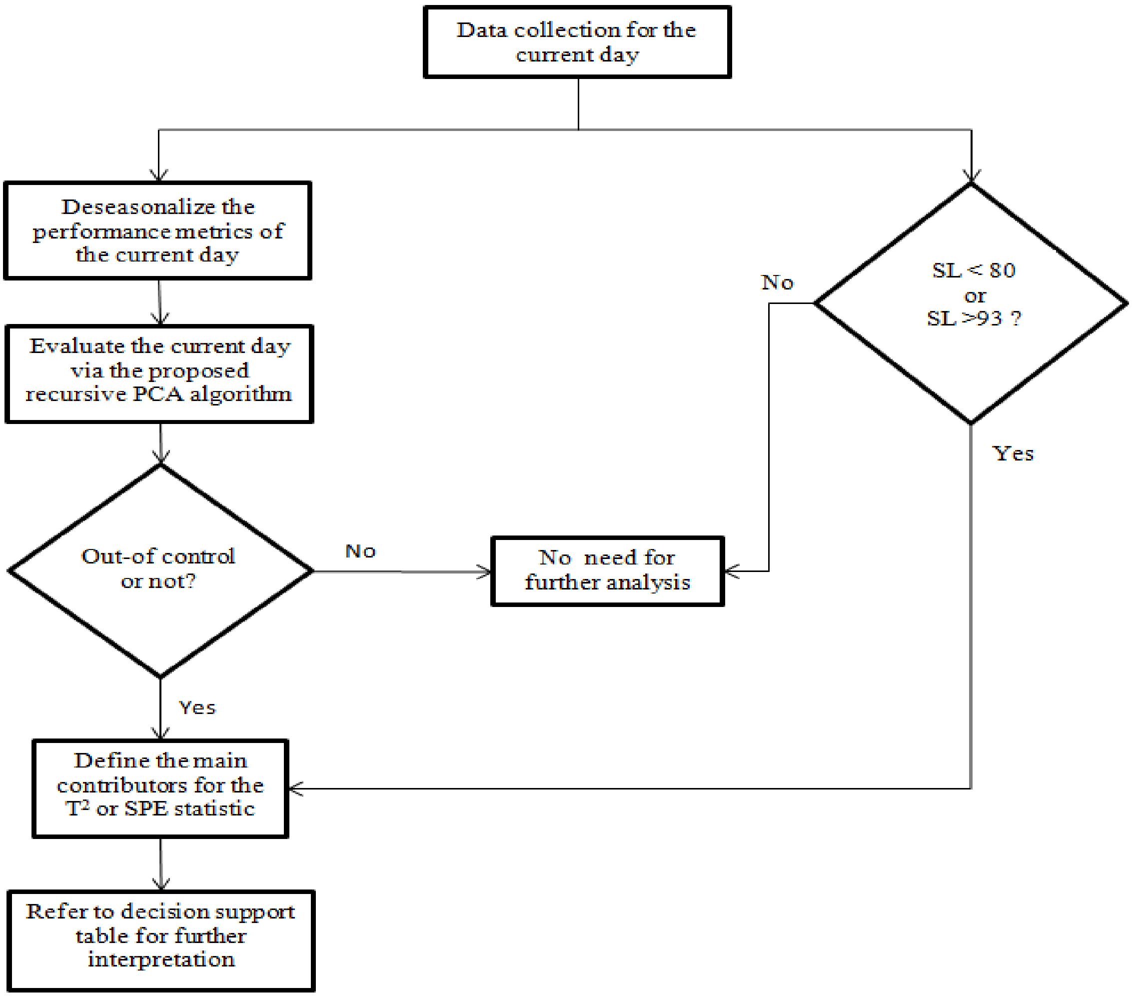

Figure 4.1. Steps of the proposed multivariate process performance evaluation method.

4.2. Monte Carlo Study of the Proposed Algorithm

The proposed algorithm's ability in detecting the wanted signals and avoiding the unwanted ones is an evident point of interest. In order to elaborate on this issue, we carried out a Monte Carlo study, using MATLAB software, to quantify these characteristics of our proposed algorithm. Firstly, we generated six multinormally distributed variables using

$$X_{generated} = \mu + C * Z \tag{4.12}$$

where $X_{generated}$ is the $p \times n$ generated multinormal data matrix, μ is p x n mean matrix, C is p x p lower triangular matrix produced by Cholesky decomposition (Equation 4.13) of the covariance matrix of $X_{generated}$, and Z is p x n matrix composed by n independent standard normal variables.

$$C * C^T = S_{X_{generated}} \tag{4.13}$$

The μ matrix was created using similar means with our six real-life variables. The C matrix was taken as the covariance matrix of an arbitrarily chosen three-month sequence of our two-year data set. So as to add the nonstationary nature of our real life data to the matrix $X_{generated}$, the values of the columns of μ were changed at every 20 observations. However, the covariance structure represented by C was kept the same for all observations.

Following this methodology, the *6 x 170* $X_{generated}$ matrix was generated. The first 70 observations were selected as the HDS. Within the other 100 observations, we injected 10 signals. That is, in 10 randomly chosen days, the values of some of the six generated values were replaced with plus or minus three standard deviation of the related mean. This signal injection step is seen in Equation 4.14.

$$X_{generated}(i, w) = \mu(i, w) \mp 3 * \sigma_i \tag{4.14}$$

where i represents the manipulated variable, w represents the day in which the signal is injected and σ_i represents the standard deviation of the manipulated variable.

Obviously, as the number of manipulated variables for a particular day increases, the probability of this day to be caught as a signal increases. For this reason, with the purpose of evaluating the power of our algorithm against variable number of manipulated variables, we have carried out this Monte-Carlo study in three parts. In the first part, 10 signals were created by manipulating only one variable for each day. In the second part, the signals were created by manipulating two variables and in the third part the signals were created by manipulating three variables for each day. These three types of signals will be called as one-variable, two-variable and three-variable signal, respectively.

For each parts of this Monte-Carlo study, the data generation and signal injection procedures were repeated through 100 iterations by changing the randomly generated Z matrix each time.

During these iterations, two different policies were followed to select the days in which the signals are injected and the variables which are manipulated. In the first policy, the days and the variables were randomly chosen within each iteration. Also, the decision about plus or minus three standard deviation in Equation 4.14 was made randomly. In the second policy, 10 particular days and related variables to be manipulated were arbitrarily chosen and kept the same for all 100 iterations. In this case, the directions of the pairwise correlations between variables were taken into account during the manipulation step employing Equation 4.14. By this way, a similarity between the types of the injected signals and signals given by the real life data was provided. The MATLAB commands of this Monte-Carlo study for the first and second policy may be found in Appendix E and F, respectively.

At each iteration, the percent of the caught signals and the percent of the unwanted signals over the total number of observations were recorded. After completing all the iterations, averages of these recorded percents were stated as the probability of detecting injected signals and the probability for unwanted signals, respectively. In Table 4.1 the result for three different types of injected signals are presented. The values in parenthesis belong to the results gathered through the second policy. As seen in the table, particularly two-variable signal cases, these numbers tend to be smaller than the others. The reason for this is that a signal involving shifts that are inconsistent with the directions of the pairwise correlations, is more probable to be detected.

Table 4.1. Results for three types of injected signals.

One-variable signals		
Φ	Pr{Detection of injected signals}	Pr{Unwanted signals}
(20,0.6)	0.58 (0.57)	0.004 (0.004)
(40,0.6)	0.78 (0.78)	0.014 (0.012)
(60,0.6)	0.81 (0.81)	0.025 (0.024)
(20,0.7)	0.55 (0.52)	0.003 (0.002)
(40,0.7)	0.72 (0.73)	0.011 (0.009)
(60,0.7)	0.80 (0.79)	0.020 (0.017)
Two-variable signals		
Φ	Pr{Detection of injected signals}	Pr{Unwanted signals}
(20,0.6)	0.87 (0.81)	0.003 (0.003)
(40,0.6)	0.94 (0.91)	0.012 (0.014)
(60,0.6)	0.95 (0.93)	0.023 (0.030)
(20,0.7)	0.84 (0.76)	0.002 (0.002)
(40,0.7)	0.92 (0.90)	0.010 (0.010)
(60,0.7)	0.95 (0.90)	0.018 (0.017)
Three-variable signals		
Φ	Pr{Detection of injected signals}	Pr{Unwanted signals}
(20,0.6)	0.96 (0.97)	0.002 (0.004)
(40,0.6)	0.98 (0.99)	0.014 (0.011)
(60,0.6)	0.98 (0.99)	0.028 (0.030)
(20,0.7)	0.94 (0.97)	0.003 (0.003)
(40,0.7)	0.99 (0.99)	0.008 (0.009)
(60,0.7)	0.99 (0.99)	0.017 (0.015)

The table designates that as the number of manipulated variables increases the probabilities of the detection of signals increases independently of the parameter couple employed by the model, as expected. Also, it is possible to argue that as WF increases detection probabilities of the injected signals tend to decrease. Moreover, for this particular generated data set, DP value of 20 performs evidently poorer than values of 40 and 60. Furthermore, it may be stated that as the length of the DP increases, both detection of injected signals and unwanted signal probabilities increase. At this point, the preference between DP values of 40 and 60 is not straightforward. However, for this particular data set, DP value 40 may be declared as the best performing one considering its fairly high *Pr{Detection of injected signals}* and lower *Pr{Unwanted signal}* compared to DP value 60.

It turns out that when the main parameters are properly selected, the proposed algorithm is capable of providing fairly acceptable results in both detecting the wanted signals and avoiding unwanted ones. Considering the absolute multivariate normality and stable covariance structure of the generated data, the algorithm is expected to perform relatively poorer in terms of *Pr{Detection of injected signals}* and *Pr{Unwanted signal}* in a real life data.

4.3. Parameter selection

This section includes the selection of DP and WF parameters. It was decided in the earlier section that a separate performance assessment process will be applied to weekdays and weekends. Thus, different DP and WF couples should be searched for these day groups.

To decide about most appropriate values of main parameters, different scenarios with different parameter couples were compared. The performances of parameter couples were evaluated based on two criteria. These are;

- Number of missed signals
- Number of unwanted signals

As a reminder, SL values higher than 93% will be considered as inefficient within our model. Thus, a missed signal represents the case where a day with an SL value lower than 80% or higher than 93%, is not declared an outlier by the algorithm. However, as it was stated in previous sections, the SL values will explicitly be traced within the performance evaluation method proposed in this study. Thus, the number of unwanted signals is of primary importance for us while deciding about the most appropriate parameter couple. The detection of unwanted signals basically depends on process knowledge. An unwanted signal means that a signaling day is analyzed to its main contributing metrics and not a noteworthy performance drift is seen.

As stated earlier the DT values will be considered as multiples of 20 and 8 for weekdays and weekends, respectively. Also, for the WF, 0.6, 0.7 and 0.8 values will be included in the trial steps. In the next two subsections different scenarios created by using different combinations of main parameters are presented. The results given by each scenario for weekdays and weekends may be seen in Table 4.2 and Table 4.3, respectively.

4.3.1. Parameter Selection for Weekdays

Firstly, we begin with the parameter couple $\Phi(20,0.6)$. The resulting *SPE* and T^2 charts are seen in Figure 4.2 and 4.3.

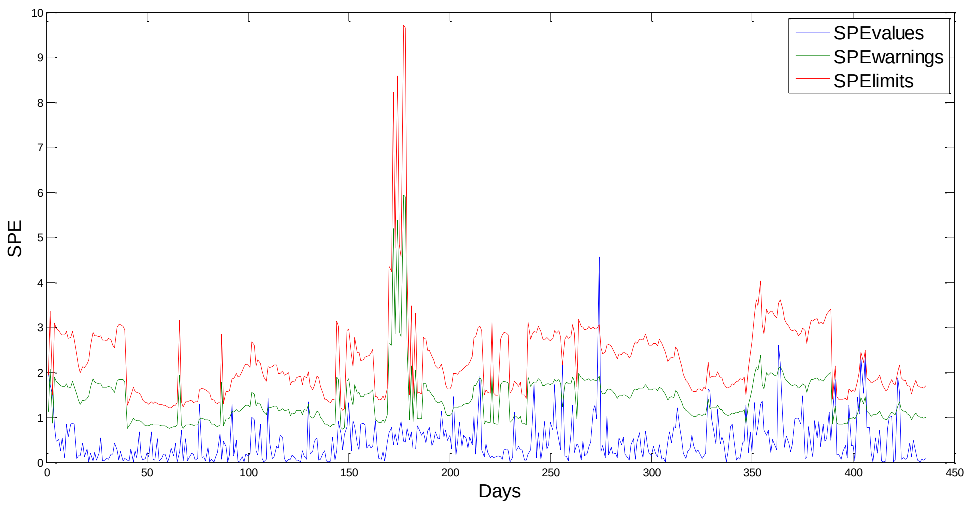

Figure 4.2. *SPE* chart for main parameters $\Phi(20,0.6)$ (Weekdays).

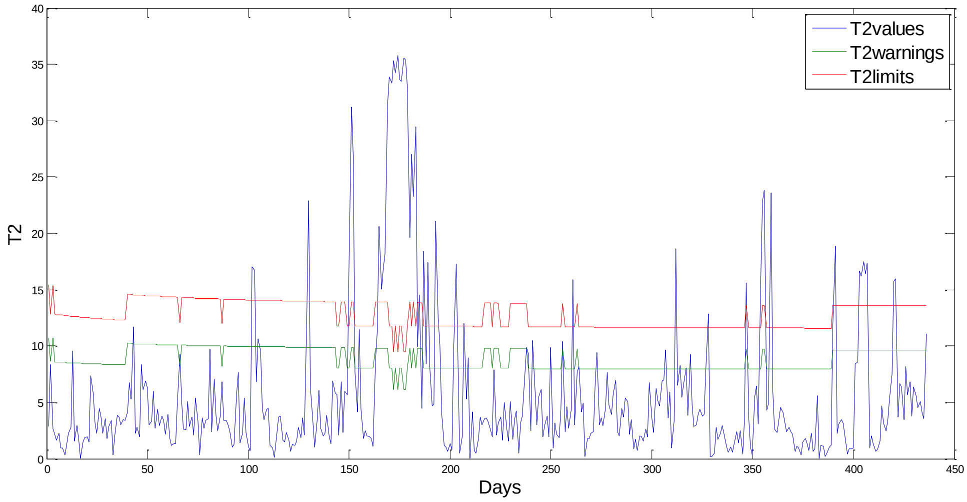

Figure 4.3. T^2 chart for main parameters Φ(20,0.6) (Weekdays).

Within this scenario, 51 signaling points were created by the model. At first sight, it is observed that a consistent increase is present for T^2 values between days 165 and 189. Although it may be considered as an inaccurate case for a control chart, as the result of interviews made with executives it was understood that the call center went through some technical problems during this period. These technical problems resulted in increased response times, decreased answering percentages and sales numbers. Detailed inspection on the model for this period showed that causes for these increases are actually the same parameters as the call center crew stated. So it was easily concluded that this peaking period in T^2 chart is in fact revealing this problematic days.

Furthermore, two other shorter periods having high increased T^2 values are seen between days 354-356 and 403-406. It was determined that these two intervals correspond to religious festivals which are general holidays for the country. Even though this is not a reason for a significant performance change, it is important for our model to catch such spikes. As stated earlier, some previously introduced recursive PCA models have a tendency to easily adapt in and miss such sudden performance drifts.

Moreover, when the other signaling points were examined, it was observed that the number of unwanted signals is just 4 which is a highly acceptable result. On the other hand, as discussed earlier, our model is expected to have a weak side in detecting some of the days with drifted service levels. With inspection of non-signaling points this weak side turned out to be clearer. However, simultaneous tracking of service level values with T^2 and *SPE* statistics remains to be a solid solution for this defect. It is still possible for someone to refer to T^2 and *SPE* contributors of a non-signaling day to find out what are the reasons of highly increased or decreased service levels.

In the next part, another scenario with a longer DP of 40 will be analyzed. The resulting *SPE* and T^2 charts for Φ(40,0.6) are seen in Figure 4.4 and 4.5. When the DP is increased to 40 having the WF fixed at 0.6, a tendency to increase is observed for both T^2 and *SPE* statistics. Also, while the limits for T^2 values remain almost the same, limits of *SPE* statistics get a lower level compared to previous scenario. This case results in a raise of signaling points. With the new parameter couple, the number of signaling points is 71.

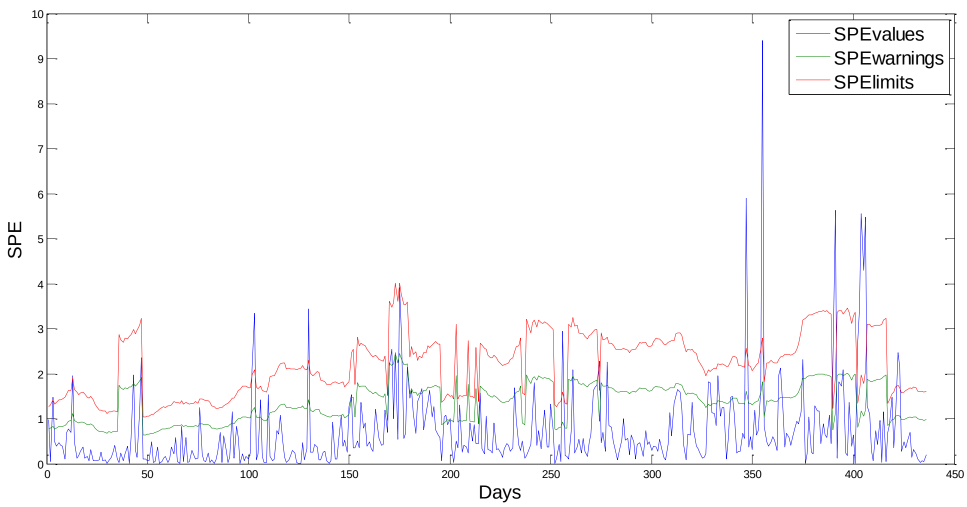

Figure 4.4. *SPE* chart for main parameters Φ(40,0.6) (Weekdays).

Close inspection of the new signaling points revealed that some of newly created signals are pointing to significant service level changes and some others are correlation related signals given by the *SPE* statistic. This increase in identification percent of wanted signals is a value-adding property for the new parameters. On the other hand, only five additional unwanted signals were created by this parameter couple. As the result, the parameter couple Φ(40,0.6) is decided to be more preferable compared to Φ(20,0.6).

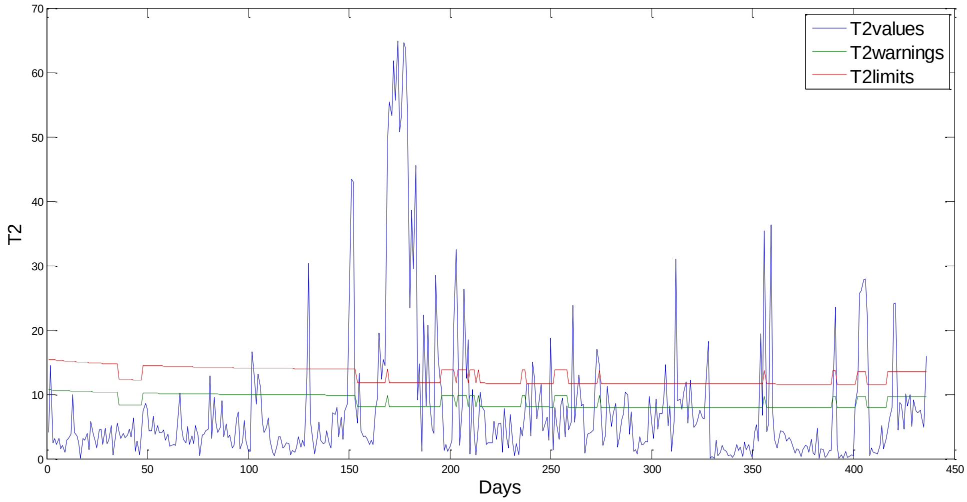

Figure 4.5. T^2 chart for main parameters Φ(40,0.6) (Weekdays).

In the next part we will increase the DP to 60 and see whether better results may be captured.

The resulting *SPE* and T^2 charts resulted by Φ(60,0.6) are illustrated in Figure 4.6 and 4.7.

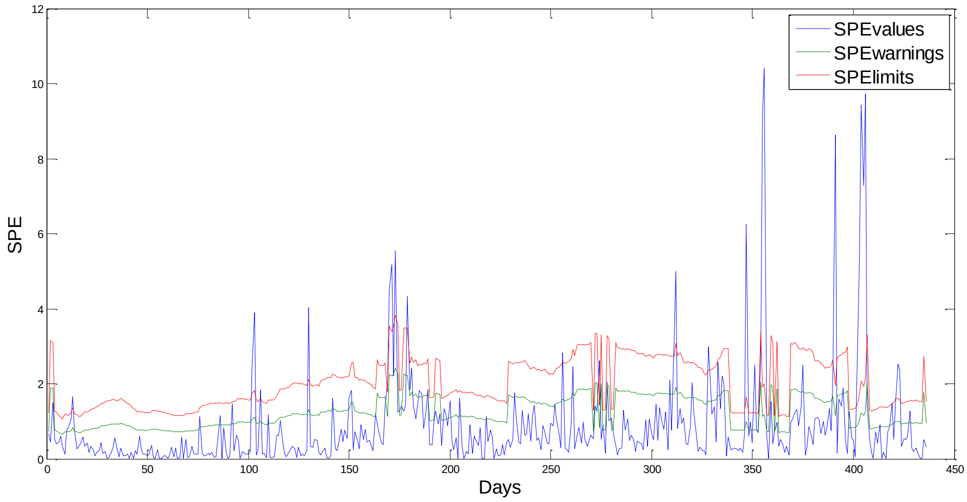

Figure 4.6. *SPE* chart for main parameters Φ(60,0.6) (Weekdays).

The total number of signaling points in this case is 96. Detailed analysis of signaling points showed that this scenario performs worse than Φ(40,0.6). Although the number of detected significant service level drifts has increased compared to both of previous scenarios, the unwanted signals have boosted.

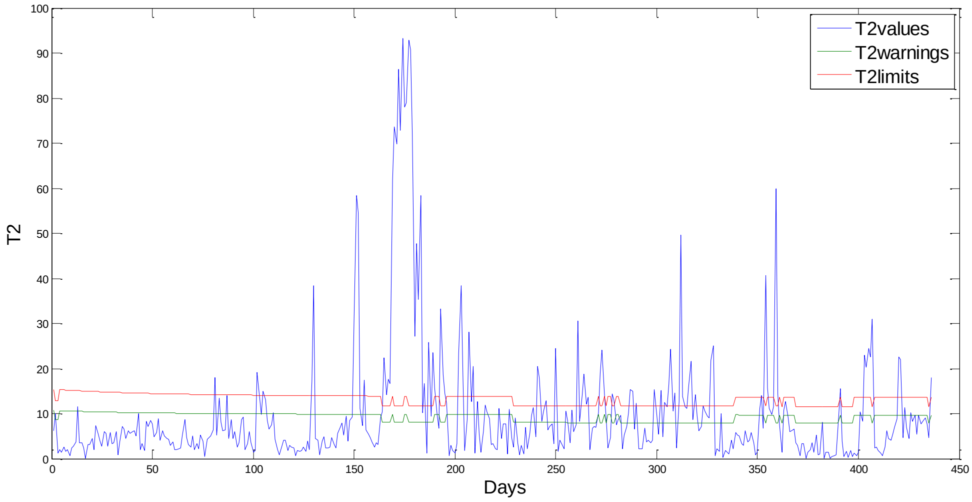

Figure 4.7. T^2 chart for main parameters Φ(60,0.6) (Weekdays).

Considering that the service levels will be externally traced, the parameter couple comprising the values of 40 for DP and 0.6 for WF remains to be the most appropriate one among those considered so far.

Obviously, assigning a higher weight on the DP is also possible. For this purpose, we will reexamine previous three cases by changing the weight of DP to 0.7. In Figure 4.8 and 4.9, the resulting *SPE* and T^2 charts for parameters Φ(20,0.7) are seen.

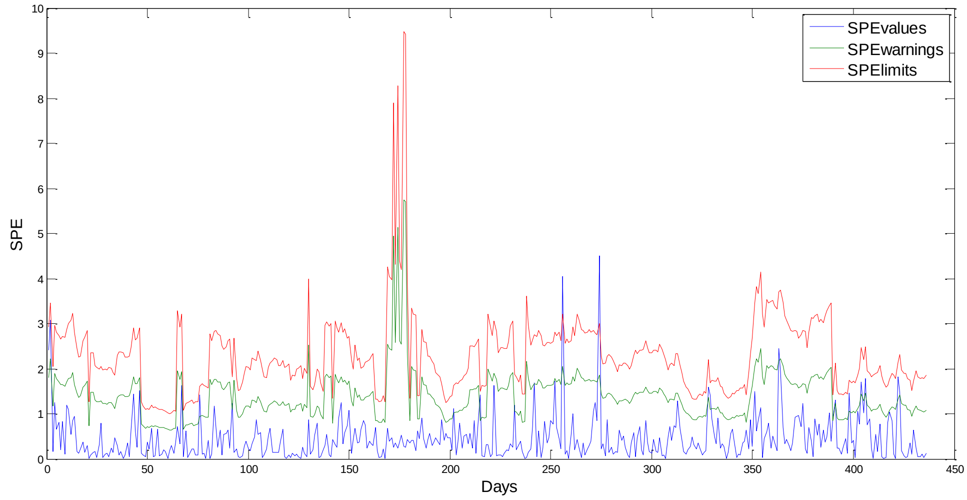

Figure 4.8. *SPE* chart for main parameters Φ(20,0.7) (Weekdays).

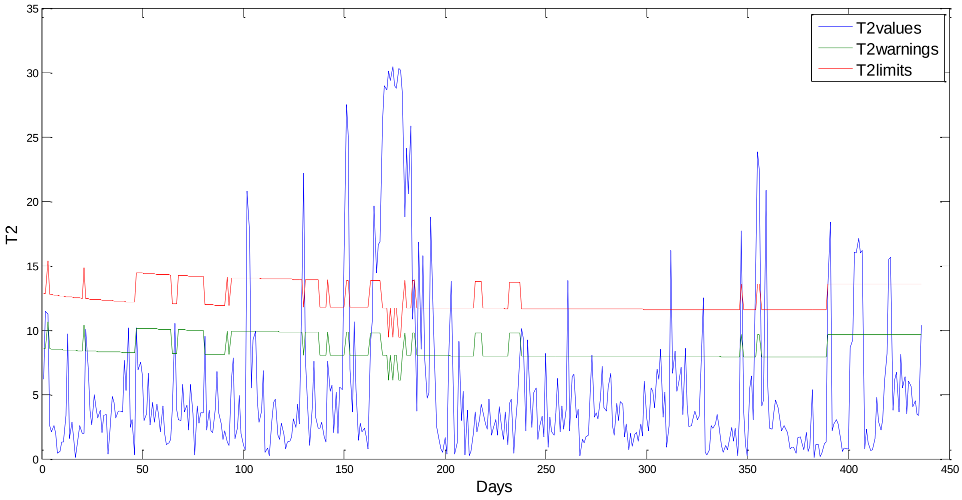

Figure 4.9. T^2 chart for main parameters Φ(20,0.7) (Weekdays).

Under normal circumstances, the number of signaling points is expected to decrease when the weight of the DP is increased. This expectation comes true for this case. Compared to the scenario Φ(20,0.6), the number of signals decreases by 7 and becomes 44. In parallel with this decrease, the number of unwanted signal goes down to 3. However, it was observed that the algorithm tends to miss some correlation related performance drifts that the parameter couples Φ(20,0.6) and Φ(40,0.6) were able to catch. For instance, days 166,167 and 168 which were designated by the model Φ(40,0.6) to suffer by faulty pop-up system interventions were not declared outliers by this new parameter couple. Assuming the potential of an increased WF for missed signals, this case leads us to conclude that the parameter couple Φ(40,0.6) is more appropriate compared to Φ(20,0.7).

Increasing the WF for DP's 40 and 60 seems to be a reasonable action, since there have been more unwanted signals for these scenarios. Figure 4.10 and 4.11 illustrate the resulting *SPE* and T^2 charts for parameters Φ(40,0.7).

With the increase in the WF parameter, the number of signaling points decreased from 71 to 68. However, close inspection created signals revealed that the number of unwanted signals still remains the same as Φ(40,0.6). As the result, the parameter couple Φ(40,0.6) is still the more appropriate one among those considered so far.

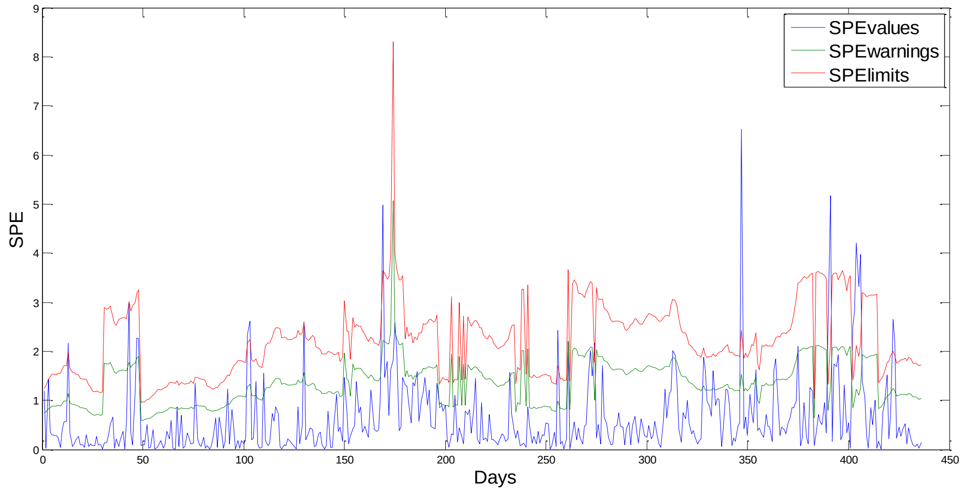

Figure 4.10. *SPE* chart for main parameters Φ(40,0.7) (Weekdays).

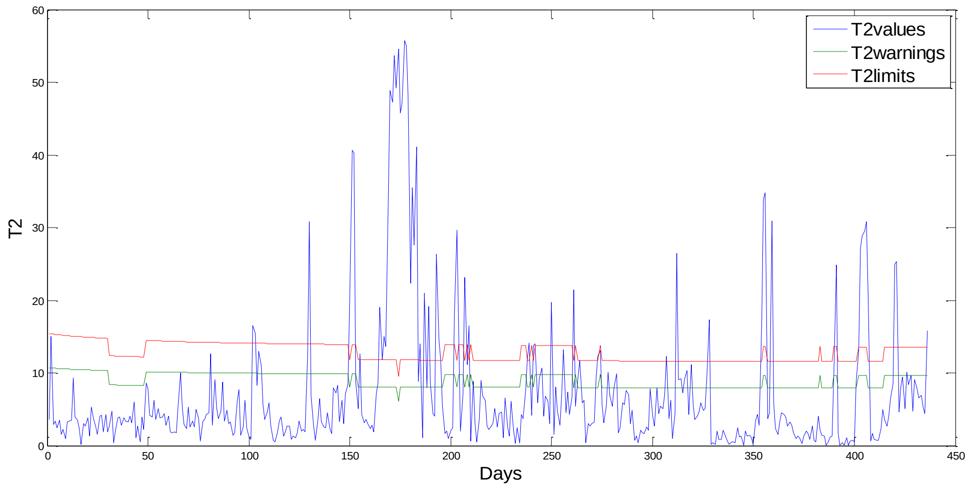

Figure 4.11. T^2 chart for main parameters Φ(40,0.7) (Weekdays).

Finally, the results given by the parameters Φ(60,0.7) will be examined. The resulting *SPE* and T^2 charts for Φ(60,0.7) are seen in Figure 4.12 and 4.13.

In this case, the number of signaling days decreases from 96 to 78. Although this decrease results in a better performance in terms of the unwanted signals, this parameter couple is still away from being the best one.

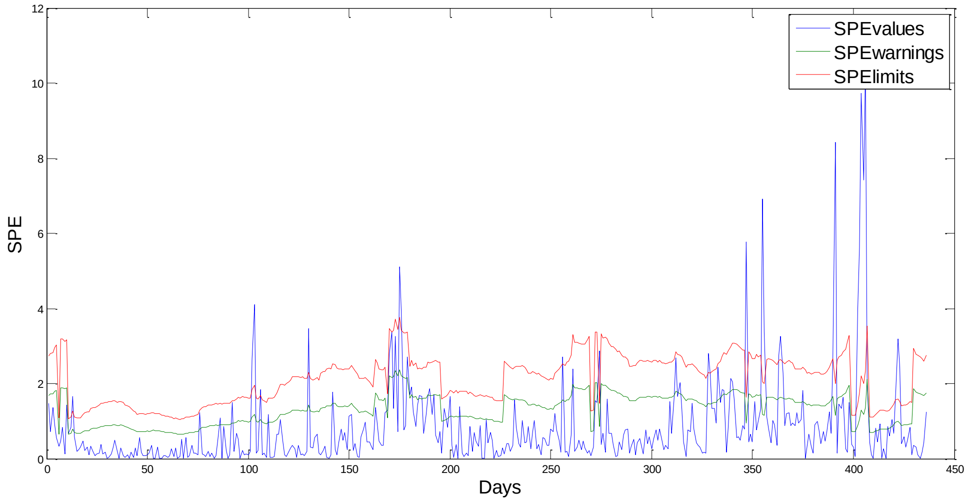

Figure 4.12. *SPE* chart for main parameters $\Phi(60,0.7)$ (Weekdays).

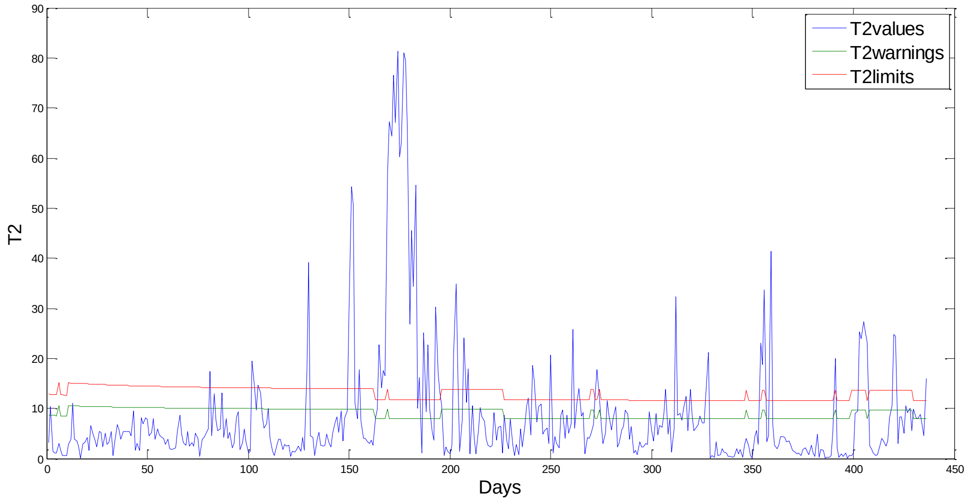

Figure 4.13. T^2 chart for main parameters $\Phi(60,0.7)$ (Weekdays).

All these three scenarios were also repeated assigning 0.8 to the WF; however no improvements were observed compared to values of 0.6 and 0.7.

The results given by different weekday model scenarios are presented in Table 4.2. Smaller number of signals given by the *SPE* statistic compared to those given by the T^2 statistic leads us to argue that the covariance structure of the weekday data set is not subject to large changes. Consequently, the parameter couple $\Phi(40,0.6)$ was decided to be used in the weekday performance evaluation model.

Table 4.2. Results of weekday model scenarios.

Φ	T^2 signals	SPE signals	Total Signals	Unwanted signals
(20,0.6)	48	3	51	4
(40,0.6)	61	18	71	9
(60,0.6)	86	27	96	18
(20,0.7)	43	1	44	3
(40,0.7)	60	17	68	9
(60,0.7)	67	26	78	15

4.3.2. Parameter Selection for Weekends

The same methodology as the one used for weekdays was followed to select the most appropriate parameters for the weekend model. As emphasized earlier, the length of the DP for weekends will be considered as the multiples of eight.

Early trial steps executed using DP of length 8 and 16 were observed to perform fairly poor in terms of the numbers of missed signals. For this reason, we begin to present parameter selection steps with the parameter couple Φ(24,0.6). The resulting *SPE* and T^2 charts for parameters Φ(24,0.6) are illustrated in Figure 4.14 and 4.15.

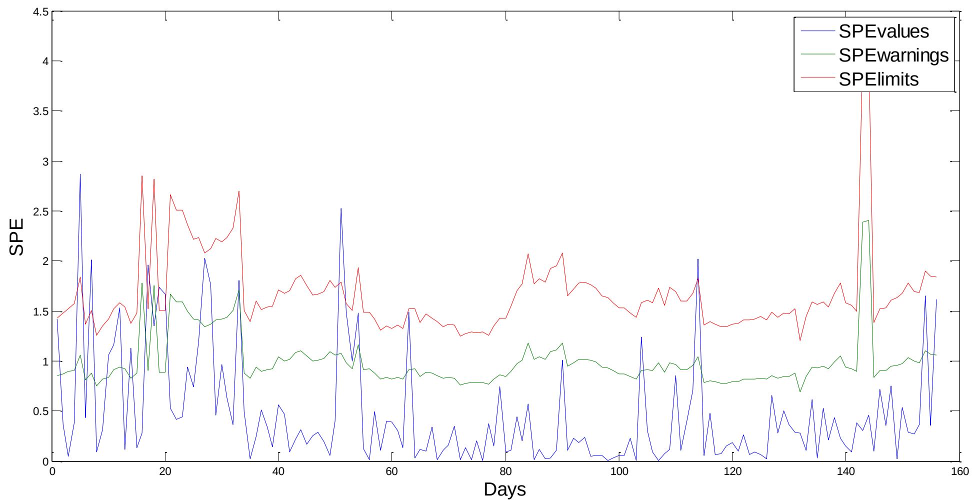

Figure 4.14. *SPE* chart for main parameters Φ(24,0.6) (Weekends).

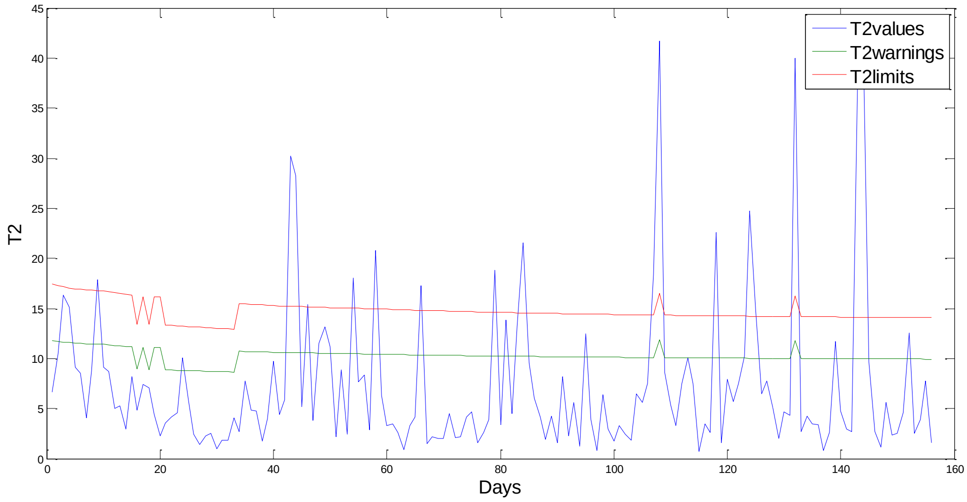

Figure 4.15. T^2 chart for main parameters Φ(24,0.6) (Weekends).

Within this scenario, 24 signals were created. Even though the model performs fairly well in terms of unwanted signals, it was observed that its missed signal performance is still subject to develop. Also, sharp elevation in the limits of *SPE* chart between days 20-30 has a potential to cause increasing missed signals. The aforementioned problematic period for the call center corresponds between days 45-60. Although this period should not necessarily be as distinct as it was for weekdays, it is expected to see relatively higher values for this period particularly in the T^2 chart. In order to check whether significant performance developments may be obtained, we increase the DP to 32 and re-execute the model. The resulting *SPE* and T^2 charts for parameters Φ(32,0.6) are seen in Figure 4.16 and 4.17.

In this case the number of signaling points has risen to 30. Analysis of new results showed that no increase is present in unwanted signals. Moreover, a general increase for T^2 values which is probably the reason for the reduction in missed signals was observed. Another gain of new parameter couple may be stated as the prevention of sharp elevation in the limits of *SPE* values. Assuming these improvements, the parameter couple Φ(32,0.6) was observed to perform best among those considered so far.

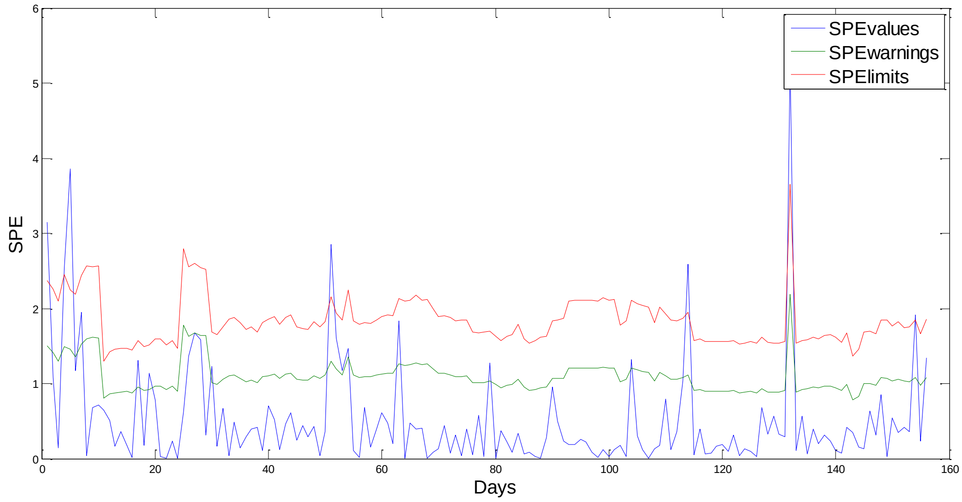

Figure 4.16. *SPE* chart for main parameters Φ(32,0.6) (Weekends).

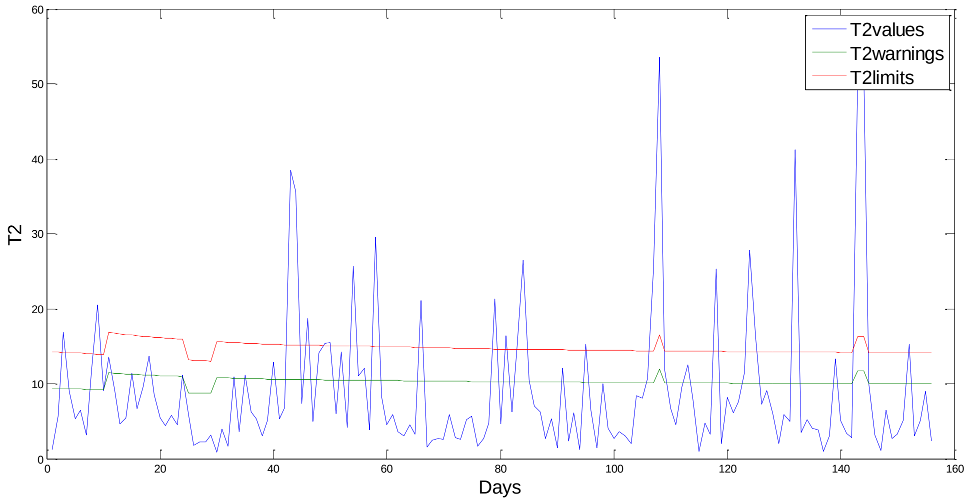

Figure 4.17. T^2 chart for main parameters Φ(32,0.6) (Weekends).

Increasing the weight of the DP is obviously an option as it was in weekday case. However, in this case, it turned out that possible improvements for the model should be searched in ways decreasing the number of missed signals. In other words, it appears that the most probable way to develop the model is to increase the number of signaling points in an appropriate range. For this reason, higher WF's are not expected to perform better for weekend model.

In Figure 4.18 and 4.19, the resulting *SPE* and T^2 charts for parameter couple Φ(32,0.7) are seen. With the increase in the WF, the number of signals went down to 25.

Since almost all the points shifting below the limits were actually wanted signals, the new parameter couple fails the test to improve model performance.

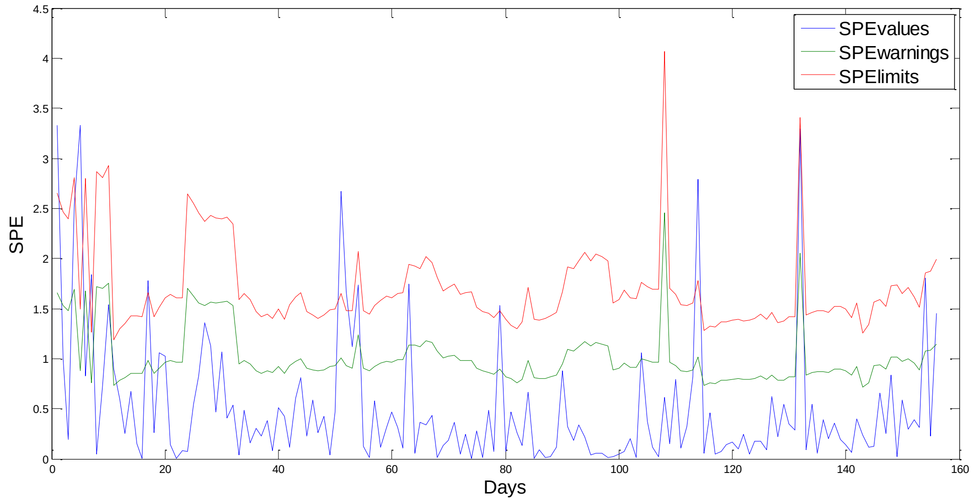

Figure 4.18. *SPE* chart for main parameters Φ(32,0.7) (Weekends).

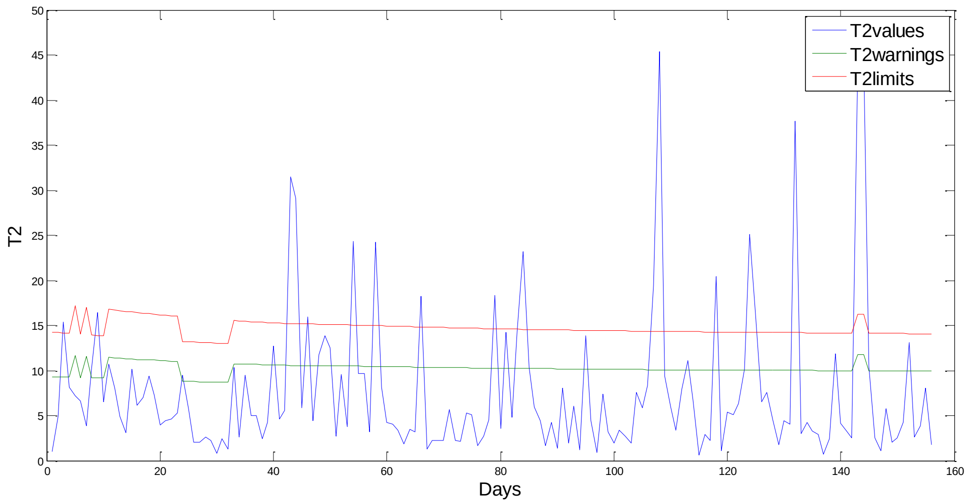

Figure 4.19. T^2 chart for main parameters Φ(32,0.7) (Weekends).

Depending on Table 4.3, as in the weekday model, considering the low percent of *SPE* signals in total signals, it may be argued that the covariance structure of the weekend data set is not subject to large changes. By expanding the HDS as required, DP lengths of 40 and 48 were also tested however; it was experienced that no performance improvements were provided by bigger DP values. Considering the results gathered through various trial steps, the parameter couple Φ(32,0.6) was decided to be used in the weekend performance evaluation model.

Table 4.3. Results of weekend model scenarios.

Φ	T^2 signals	SPE signals	Total Signals	Unwanted signals
(24,0.6)	17	7	24	2
(32,0.6)	24	7	30	2
(24,0.7)	15	5	20	0
(32,0.7)	18	8	25	1

In Table 4.4 the details of the models constructed depending on the parameter selections for weekdays and weekends, are presented.

Table 4.4. Details of constructed models.

	Weekday model	Weekend model
Length of the HDS	68	39
DT	40	32
WF	0.6	0.6
Number of days evaluated	456	156
Number of Signals	71	30
Number of unwanted signals	9	2

4.4. Detailed Analysis of Selected Scenarios and Decision Support Table

Having determined the most appropriate parameters for weekday and weekend models, in-depth examination of characteristic signaling points is the topic of this section.

At the end of the section, a decision support table involving commonly encountered contributing variables and their practical interpretations will be given.

SPE and T^2charts resulted by the selected scenarios for weekdays and weekends are seen in Figure 4.4, 4.5, and Figure 4.16, 4.17, respectively. T^2 value for 102 was observed to be out of control limits. Below are the contribution plots for increased T^2 value. It is important to note that the absolute values of these numbers should be considered while assessing the magnitudes of contributions made by each variable.

Additionally, it is possible for one to hesitate about considering a particular performance metric as a main contributor. In our study, for most of the signals, the main contributors are clearly visible in contribution plots. However, in cases where this judgment is not so straightforward, it should be noted that these are individual contributions of variables composing an observation which has already created a signal. Therefore, rather than searching for a strict statistical significance, it is wiser to depend on process knowledge and experience to make this decision. (AlGhazzawi and Lennox, 2008) handled this problem by simply referring to the individual data plots of the variables of interest. With a similar point of view, we recommend a guideline for this problem in our study. After having necessary discussion with call center executives,

$$\textit{Weighted Mean} \mp 1{,}5 * \textit{Weighted Standard Deviation}$$

was determined as a reference range for this question. At this point, if there is an hesitation about the main contributors, it is advised to name a performance metric as a main contributor when its deseasonalized original value falls outside this range.

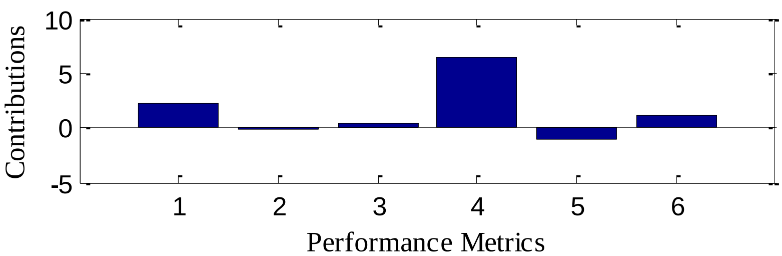

Figure 4.20 Contributions for weekday 102.

As seen in Figure 4.20, the out-of-control T^2 value in weekday 102 is caused by the fourth variable which is *CpA*. Obviously, two cases are possible for a contributing variable; first one is to be higher and the second one is to be lower than it should be. To decide about whether this contribution is caused by a high or a low value, the comparison between the current value and the weighted average is a reference point. It was determined that a low call per agent value is the reasons for the signal. Additionally, the SL values is seen to be 94% which is higher than 93% efficiency threshold. So, it may be concluded that the number of agents recruited for this day is higher than it should be. This case points out that the call center executives followed an inefficient resource allocation policy at this day.

When we look at the weekday 130, the only contributing variable is seen to be the first metric which is *AP*.

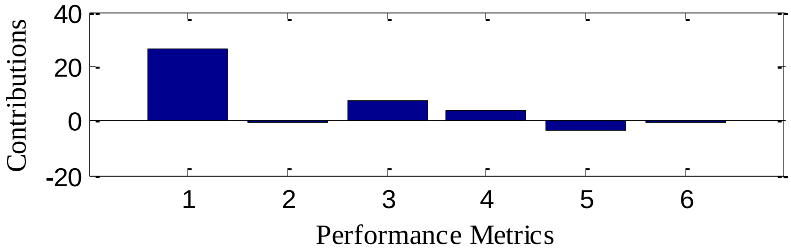

Figure 4.21. Contributions for weekday 130.

A decrease in this variable seems to be the reasons for this signal. It is also observed that the related SL (80%) is also at the lower bound of acceptable area. It turns out that the average waiting time before call loss values are also supposed to be examined so as to be able to make a clearer comment for such a case.

In weekday 150, low *SpC* and $\overline{WT}$ values are the main contributors. Since it is not possible to talk about a causal link between these two metrics they should be interpreted separately. The decrease in *SpC* simply highlights a poor sale performance of the agents. On the other hand, the decrease in the $\overline{WT}$ may have several reasons. In any cases, this variability in working times may turn out to be a threat to decrease SL value. However, the SL value for this particular day is 83%. This value should be considered as the result of a successful in-day resource allocation.

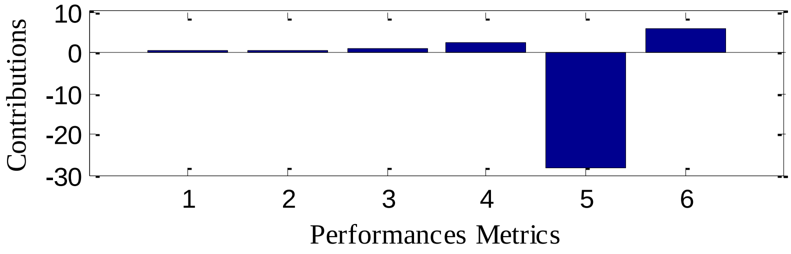

Figure 4.22. Contributions for weekday 150.

Between weekdays 169-189, as the result of technological problems suffered by the call center, *AP*, $\overline{RT}$, and *SpC* metrics are the most commonly faced contributing variables. Within this period, decreased answer percents and sales numbers are the main contributors together with increased average response times. Since this is an exceptional period, no much inference and detailed analysis is needed about the reasons of signaling days.

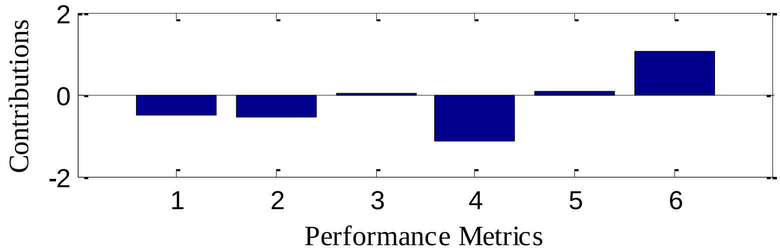

Figure 4.23. Contributions for weekday 256.

In weekday 256, the , *CpA* and $\overline{WT}$ are seen to be the main contributors for an out-of-control *SPE* value. Highly decreased *CpA* values designates to a resources allocation problem. As stated earlier, detecting efficiency related problems is an important intention of our model. The 95% SL value for the day supports that not an efficient forecasting and resource allocation policy was followed. Also, an it seems that increased $\overline{WT}$ values have contributed to the catch an inefficiently high SL. Further investigation on the day reveals that it is an official holiday in the country. At this point what should be questioned is that whether it was needed to have such a big number of working agents costing a big amount for the company or it was better to have a smaller number of agents with a still acceptable service level? The answer for this question should be obviously “Yes” for a decision maker having an efficiency based performance perception.

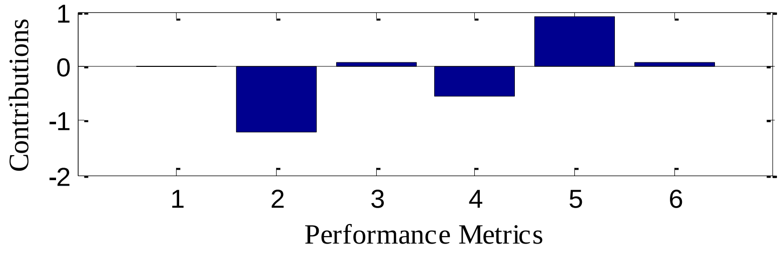

Figure 4.24. Contributions for weekend 114.

In weekend 114, $\overline{ACD}$ and SpC metrics are the main contributors of the signal given by the SPE statistic. The noteworthy side of this signal is that although a relatively big contribution is made by $\overline{ACD}$, it is not a magnitude but correlation reasoned contribution. That is, even though the SpC number increases, the $\overline{ACD}$ remains in an ordinary range. This case reveals that a high level sale performance was shown by the agents.

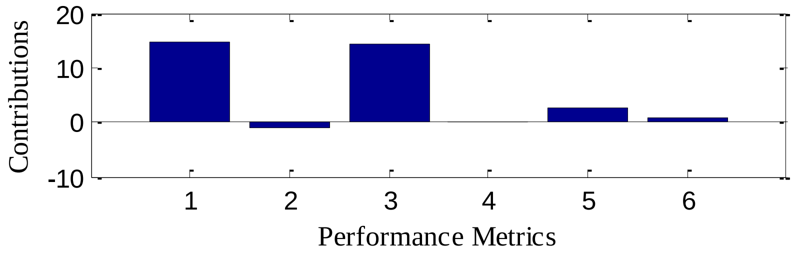

Figure 4.25. Contributions for weekday 261.

In weekday 261, the contribution plot highlights two metrics as being source of the violation, AP and $\overline{RT}$. Investigation of these two variables indicates that low answering percents and increased response times have caused this performance drift. In such a case, a decrease in service levels is almost unavoidable. The 73% SL value is verifying this expectation. At first sight, it seems that a bad resource allocation caused by a forecasting problem is the reason for the drift. However, the most probable sign of bad resource allocation is a significant contribution made by variable CpA. Such a case is not present in this day. For this reason, the performance drift should be interpreted to be caused by an unexpected workload. Since the uncertainty and randomness is the nature of forecasting business, such a performance fluctuation has an acceptable origin in a decision maker's point of view.

On the other hand, as stated earlier, manipulation of pop-up screens is the main preventive action in call center management against increased queues. However, low contributions of $\overline{ACD}$ and SpC metrics prove that no proactive actions were taken to prevent decreasing service levels. As the result, it is possible to conclude that a poor pop-up management was conducted by team leaders.

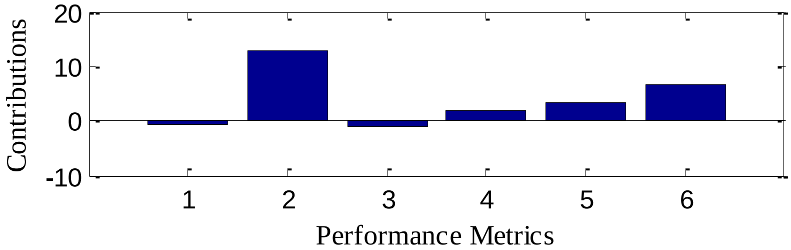

Figure 4.26. Contributions for weekend 46.

As mentioned earlier, although they are closely related, it is possible to observe some incoherent behaviors of service level and average response time statistics. A clear appearance of such a case is present in weekend 46. The contributions made by each metric are seen above. It was examined that low $\overline{ACD}$ and SpC values are two of the three main contributors for the signal. Since not a significant contribution is given by the third variable, it seems that an unnecessary sales closing action was conducted by the executives. However, explicit examination of SL values shows that although the $\overline{RT}$ has almost no contribution to the signal, the SL value of 80% is right on the lower bound of the acceptable area. As the result, the preventive pop-up system manipulation of the executives turns out to be a clever action to force the SL value stay over the 80% limit.

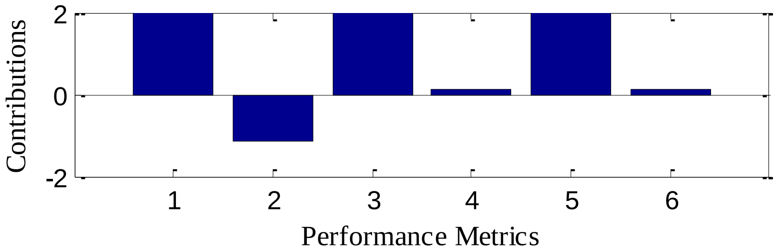

Figure 4.27. Contributions for weekday 312.

Another characteristic signal where AP, $\overline{ACD}$, $\overline{RT}$ and SpC metrics are all big contributors, is given by T^2 chart for weekday 312. Several types of contribution combinations are possible when the contributors are evolved in this way; however two of them are the most common ones. The first case is the consequence of a compulsory pop-up system manipulation where increasing $\overline{RT}$'s and decreasing *AP*'s are tried to be balanced by preventive sale closing actions. Evidently, such as day should be interpreted as a successful pop-up management suffering of unexpected workloads. On the other hand, the second case is faced when unnecessary sale closing actions resulted in inefficiently decreased $\overline{RT}$. Obviously, the SL value is the distinguishing factor among these two cases. Weekdays 420 and 421 are good examples for both cases. Having a low SL value (69%), weekday 312 pertains to the first case where a compulsory pop-up system manipulation is of issue. On the contrary, concerning the SL values of 94 and 90%, it may be concluded that an unnecessary sales interventions were effectuated by the team leaders at weekdays 420 and 421.

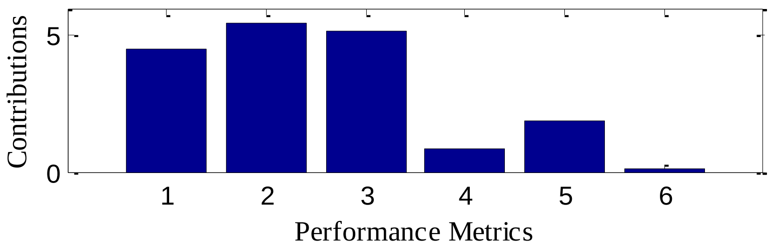

Figure 4.28. Contributions for weekday 420.

Consistent rise in T^2 values between weekdays 403-406 was determined to reflect an efficiency problem. As seen in Figure 4.28 all four days have similar structures in terms of contributing variables. The *AP*, $\overline{ACD}$, $\overline{RT}$, *CpA* are the main contributors for all days. Further investigation on this period demonstrated that it is an official holiday for the country. As it is the case for previously analyzed official holidays, an inefficient resource allocation is again the reason for the signals. High *AP*, low $\overline{RT}$, and low *CpA* values designates to this inefficiency problem. This inference is supported by SL values of 95%, 91%, 96%, and 94% for four consecutive days. Except the day 404 the SL values are above the 93% threshold. Also, the contributions made by $\overline{ACD}$ values should separately be evaluated than other performance metrics. Decreased *ACD* times seems to be an uncontrollable variation for the call center.

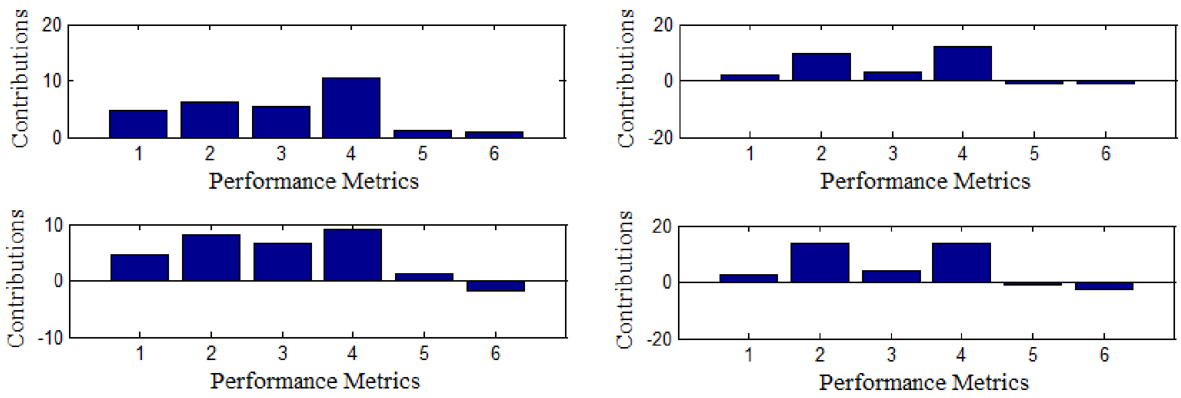

Figure 4.29. Contributions for weekdays 403-406.

As such analyses about the contribution structures and reasons of signaling points may be expanded, a decision support table that involves the most commonly encountered types of signals and related inferences, will be introduced in Table 4.6. The table makes it easier for a decision maker to analyze the reasons of a signal.

The decision support table was constructed by analyzing the signals of both weekday and weekend models. During these analyses, we had necessary discussions with call center executives. Depending on their consultations, the most likely reasons for drifted performances were defined as in Table 4.6. If one encounters with a type of signal that may be classified by a combination of the signal types included in this table, each table row should be referred individually to evaluate the signal. It should be noted that the statements in this table are just probable predictions about the reasons of drifted metrics and they are subject to change depending on process knowledge. Obviously, this table should be more detailed and precise to become a comprehensive tool. A more intensive collaboration of call center executives is of primary importance for the improvement of this table.

The weighted mean and weighted standard deviation values for the related variables are designated within the table as $W.M_{(i)}$ and $W.Std_{(i)}$, respectively. A type of signal composed of any combination of the types of signals included in this table, should be evaluated by separately referring to each row. In this table, the six performance metrics will be represented as X_i. Table 4.5 represents the counterparts of each performance metric.

Table 4.5. Counterparts of performance metrics.

$X_1 = AP$
$X_2 = \overline{ACD}$
$X_3 = \overline{RT}$
$X_4 = CpA$
$X_5 = SpC$
$X_6 = \overline{WT}$

In Figure 4.30 the steps of the proposed performance assessment system are illustrated. It is essential to note that although the developed performance monitoring tool and decision support table are able to identify performance drifts and related reasons, it is evident that process knowledge and experience has an essential role in analyzing the results and validating the inferences.

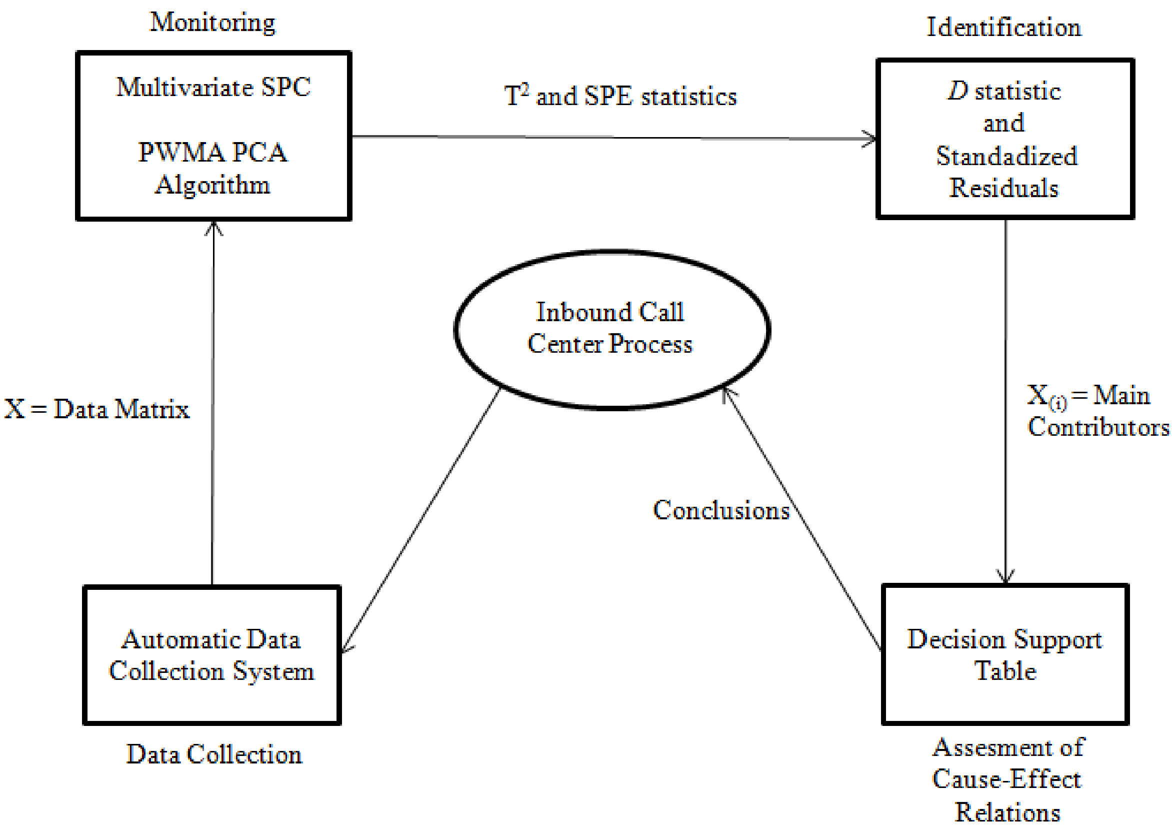

Figure 4.30. Proposed performance assessment system.

Table 4.6. Decision Support Table.

<table>
<tr><th>Signal Type</th><th>Main Contributors</th><th>Signaling Statistic</th><th>Preliminary Action (if necessary)</th><th colspan="4">Most likely reasons for drifted performance metrics</th></tr>
<tr><td rowspan="3">1</td><td rowspan="3">X_2, X_5</td><td rowspan="3">T^2</td><td rowspan="3">Check SL</td><td>SL > 85</td><td>80<SL<85</td><td>SL>85</td><td>SL<80</td></tr>
<tr><td colspan="2">$X_2 < W.M_{(2)}$ and $X_5 < W.M_{(5)}$</td><td colspan="2">$X_2 > W.M_{(2)}$ and $X_5 > W.M_{(5)}$</td></tr>
<tr><td>Poor pop-up management. Improper sale closing decision</td><td>Good pop-up management.Necessary preventive actions were taken</td><td>High sale performance</td><td>Poor pop-up management. Necessary preventive actions were not taken</td></tr>
<tr><td rowspan="2">2</td><td rowspan="2">X_2 , X_5</td><td rowspan="2">SPE</td><td rowspan="2">---</td><td>$X_2 > W.M_{(2)} + 1.5*W.Std_{(2)}$ and $X_5 < W.M_{(5)} + 1.5*W.Std_{(5)}$</td><td colspan="2">$X_2 < W.M_{(2)} + 1.5*W.Std_{(2)}$ and $X_5 > W.M_{(5)} + 1.5*W.Std_{(5)}$</td><td>$X_2 > W.M_{(2)} + 1.5*W.Std_{(2)}$ and $X_5 > W.M_{(5)} + 1.5*W.Std_{(5)}$</td></tr>
<tr><td>Inefficient sale performance</td><td colspan="2">Efficient sale performance</td><td>High sale performance</td></tr>
<tr><td rowspan="2">3</td><td rowspan="2">X_1, X_3, X_2, X_5</td><td rowspan="2">T^2 or SPE</td><td rowspan="2">Check SL</td><td colspan="2">SL > 85</td><td colspan="2">SL < 85</td></tr>
<tr><td colspan="2">Poor pop-up management. Improper sale closing decision</td><td colspan="2">Unexpected workload and good pop-up management</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">X_1, X_3</td><td rowspan="2">T^2</td><td rowspan="2">---</td><td colspan="2">$X_1 < W.M_{(1)}$ and $X_3 > W.M_{(3)}$</td><td colspan="2">$X_1 > W.M_{(1)}$ and $X_3 < W.M_{(3)}$</td></tr>
<tr><td colspan="2">Unexpectedly high workload and poor pop-up management</td><td colspan="2">Low workload</td></tr>
<tr><td>5</td><td>X_1, X_3, X_4</td><td>T^2</td><td>Check SL</td><td colspan="2">SL > 93 --- Inefficient resource allocation</td><td colspan="2">SL < 80 --- Unsufficient resource allocation</td></tr>
<tr><td rowspan="2">6</td><td rowspan="2">X_4, X_6</td><td rowspan="2">T^2 or SPE</td><td rowspan="2">---</td><td colspan="2">$X_4 < W.M_{(4)}$ and $X_6 < W.M_{(6)}$</td><td colspan="2">$X_4 > W.M_{(4)}$ and $X_6 > W.M_{(6)}$</td></tr>
<tr><td colspan="2">High potential for low SL. Revise resource allocation</td><td colspan="2">Inefficient resource allocation.</td></tr>
<tr><td>7</td><td>X_2</td><td>SPE</td><td>Check X_5</td><td colspan="4">Refer to row 2 for further interpretation</td></tr>
<tr><td>8</td><td>X_5</td><td>SPE</td><td>Check X_2</td><td colspan="4">Refer to row 2 for further interpretation</td></tr>
</table>

Table 4.7. Decision Support Table (cont.)

<table>
<tr><td>9</td><td>X_1</td><td>T^2</td><td>---</td><td colspan="3">Uncontrollable signal. Probably decreased average waiting time before call loss</td></tr>
<tr><td>10</td><td>X_2</td><td>T^2</td><td>---</td><td colspan="3">Uncontrollable signal. Low or high talking times</td></tr>
<tr><td rowspan="2">11</td><td rowspan="2">X_4</td><td rowspan="2">T^2 or SPE</td><td rowspan="2">Check SL</td><td colspan="2">$SL > 93$</td><td>$SL < 80$</td></tr>
<tr><td colspan="2">Inefficient resource allocation.</td><td>Unsufficient resource allocation. Probably faulty forecast.</td></tr>
<tr><td rowspan="2">12</td><td rowspan="2">X_5</td><td rowspan="2">T^2</td><td rowspan="2">---</td><td colspan="2">$X_5 < W.M_{(5)}$</td><td>$X_5 > W.M_{(5)}$</td></tr>
<tr><td colspan="2">Poor sale performance</td><td>High sale performance</td></tr>
<tr><td rowspan="2">13</td><td rowspan="2">X_6</td><td rowspan="2">T^2 or SPE</td><td rowspan="2">Check SL</td><td>$80 < SL < 93$</td><td>$SL > 93$</td><td>$SL < 80$</td></tr>
<tr><td>Successful resource allocation to deal with variable working times</td><td>Inefficient resource allocation</td><td>Unsufficient resource allocation. Unable to deal with variable working times</td></tr>
</table>

5. CONCLUSIONS AND FUTURE WORKS

This work presented the development of a performance monitoring system for an inbound call center process. The main aim of the study was to come up with an efficiency based performance assessment tool that is able to reveal performance drifts and related causes. For this purpose, a new approach was created for inbound process performance evaluation. This new approach translates the conventional univariate performance assessment method into a multivariate performance evaluation model. Key characteristics of the proposed model are listed in four elements:

- Efficiency viewpoint and consideration of causal links between performance metrics,
- Simultaneous tracking of service level and newly proposed performance metrics,
- Separate performance assessment models for weekdays and weekends,
- A recursive PCA algorithm which is capable of tracing time-varying nature of call center dynamics.

The proposed Periodic Weighted Moving Average Principal Component Analysis (PWMAPCA) algorithm is an alternative to the recursive PCA algorithms introduced by (Li *et al*., 2000), (Lane *et al*., 2003) and (AlGhazzawi and Lennox, 2009). In the algorithms proposed by (Li *et al*., 2000) and (Lane *et al*., 2003), forgetting factors are used to adapt the model to inherent dynamic flow of the process. (Wang *et al*., 2003) and (Sandoz, 2003) proved that employing forgetting factors makes an algorithm prone to miss sudden drifts in the process. On the other hand, in the algorithm constructed by (AlGhazzawi and Lennox, 2009), the forgetting factor is replaced by a stable-size moving window approach. In their model, at each time a new observation becomes available, a new PCA is executed depending on only a particular number of most recent observations. Our PWMAPCA algorithm is a modified version of the one proposed by (AlGhazzawi and Lennox, 2009). It is based on the update of the mean, standard deviation and the covariance structure depending on the weighted averages of a particular number of most recent observations and the previous days. While the number of most recent observations whose weights are increased, is represented by DP, the weight of this period is represented by WF (>0.5).

A Monte Carlo study was carried out, using MATLAB software, to test the proposed algorithm's ability in detecting the wanted signals and avoiding the unwanted ones. This Monte Carlo study showed that when the main parameters are properly selected, the algorithm is able to perform fairly well in both detecting the wanted signals and avoiding unwanted ones.

Six performance metrics for the call center process were defined in a way to strengthen inferential abilities of the model. Two years of data were collected. Because of strong seasonality between weekdays and weekends, it was decided to implement separate performance assessment models for each day group. Consequently, two different parameter couples were determined and the same algorithms employing different parameter couples were used in the performance assessment process of weekdays and weekends.

As the result of the study, a PCA based performance monitoring tool tracing the daily performances of the inbound department of a call center was created. The tool is well capable of keeping up with the nonstationary structure of the process. A collaborative approach involving simultaneous tracking of service levels and results gathered from the monitoring tool was proposed in order to minimize the number of missed signals.

Two years of data were analyzed by the proposed model. The historical data set of weekdays was selected from within the first four months, while the first six months was used for the same purpose for weekends. After making necessary adjustments for seasonality and autocorrelation, the data were passed through the model. Considering the weekdays and weekends as separate day groups, 436 weekdays and 156 weekends were examined. 71 days for weekdays and 30 days for weekends were observed to have significant performance changes. Successful and poor pop-up system managements, efficient or inefficient resources allocations and high or low sales performances of agents may be classified as the main headings of detected performance drifts.

A decision support table involving the most probable types of performance drifts and practical interpretations about their reasons was given at the end of the study. It was also noted that depending on the process knowledge and experience, interpretations of some signals are subject to differ from the ones stated in this table.

Also, the study showed that the PCA method, which has been mostly used for industrial processes, may efficiently be adapted into a service system performance evaluation work. Although, it was implemented in a call center process in this work, the proposed tool is evidently available to be used for various service systems by properly defining the performance metrics and parameter couples. Additionally, in this study, the performance evaluation was conducted based on daily values. By reducing this time scale, and collecting the data for smaller time sequences, it could be possible to construct a preventive performance monitoring system. However, it should be noted, when the time scale is reduced, some modifications on the proposed algorithm against probably higher autocorrelation levels and time constraint, may be required.

As stated in previous chapters, some performance metrics could not be added into the model due to technological constraints. Inclusion of such value-adding variables would definitely augment the performance and inferential abilities of the model. Moreover, another open-to-progress point of the method is the selection of two main parameters. As explained in the parameter selection part, the parameters were selected in a comparative way depending on results given by different parameter couples. Development of a more strict and robust methodology for the determination of most appropriate parameters should be the main future research area on this work. Furthermore, the decision support table provided at the end of the study is obviously subject to be developed. With a stronger collaboration of call center executives, this table may be turned into a more detailed and precise decision support tool.

APPENDIX A: AUTOCORRELATION FUNCTIONS OF TWO-YEAR DATA SET

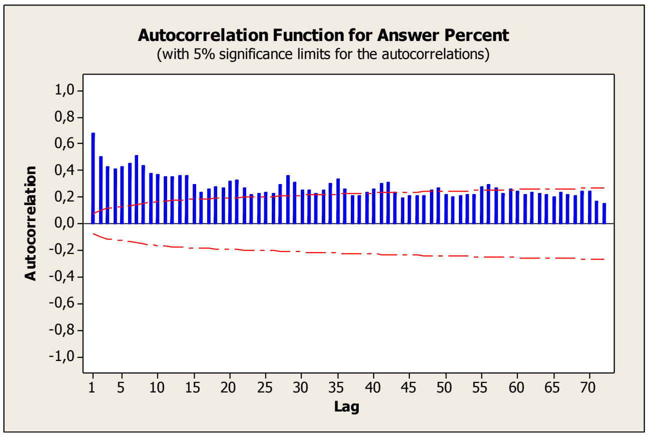

Figure A.1. Autocorrelation function for *AP*.

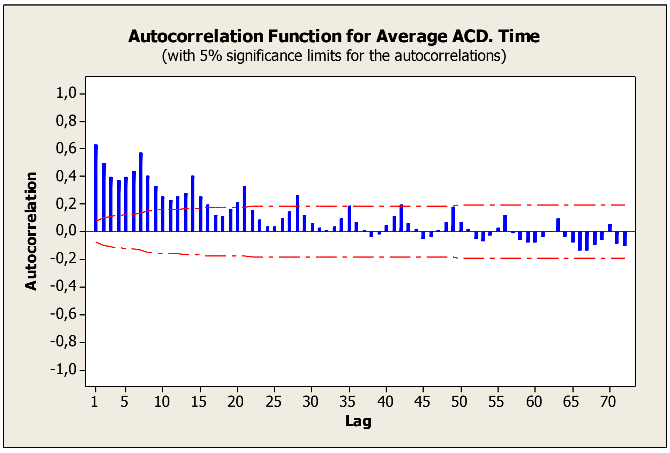

Figure A.2. Autocorrelation function for $\overline{ACD}$.

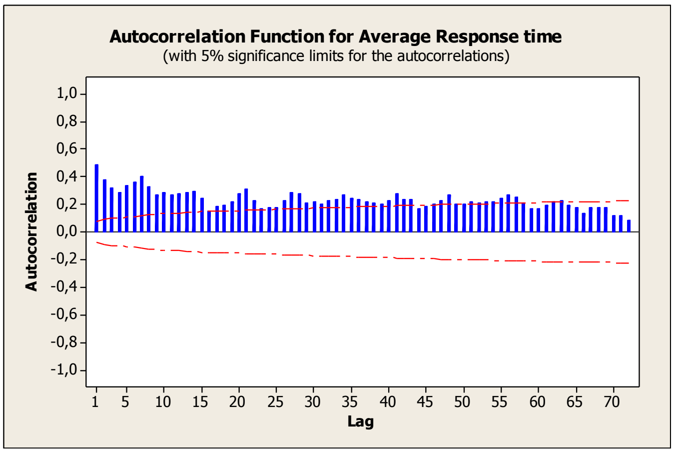

Figure A.3. Autocorrelation function for $\overline{RT}$.

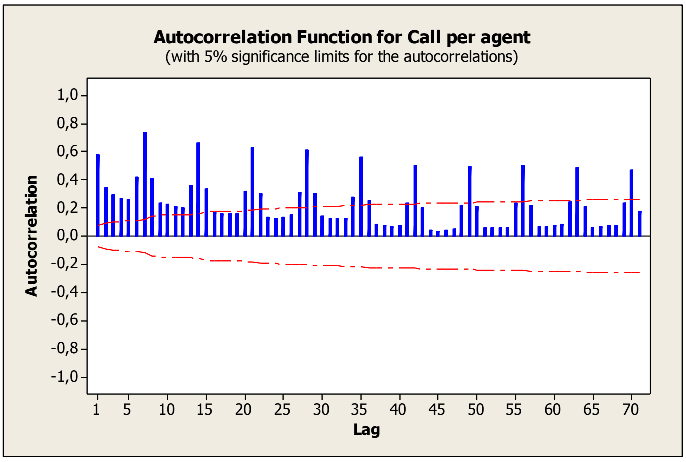

Figure A.4. Autocorrelation function for *CpA*.

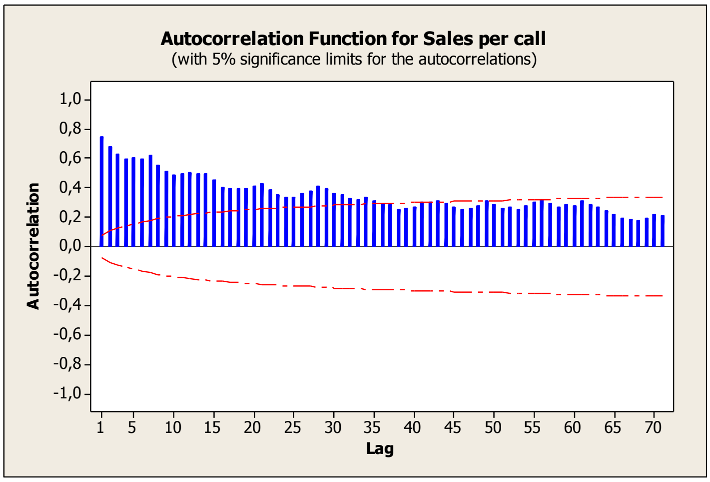

Figure A.5. Autocorrelation function for *SpC*.

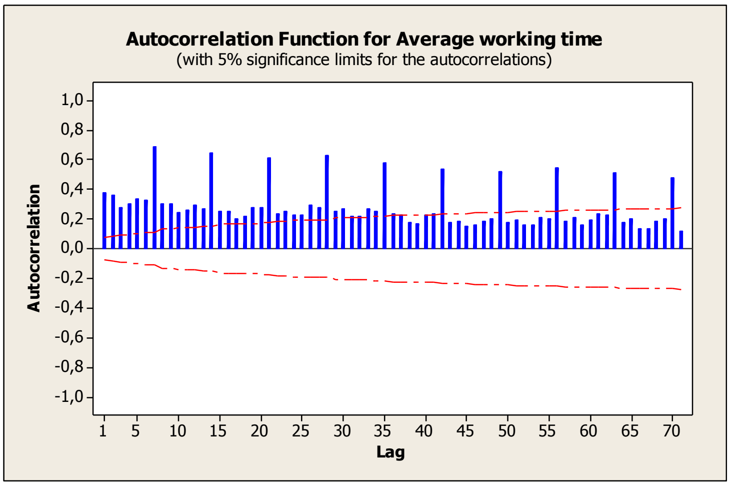

Figure A.6. Autocorrelation function for $\overline{WT}$.

APPENDIX B: AUTOCORRELATION FUNCTIONS OF WEEKDAYS AND WEEKENDS (ROW DATA)

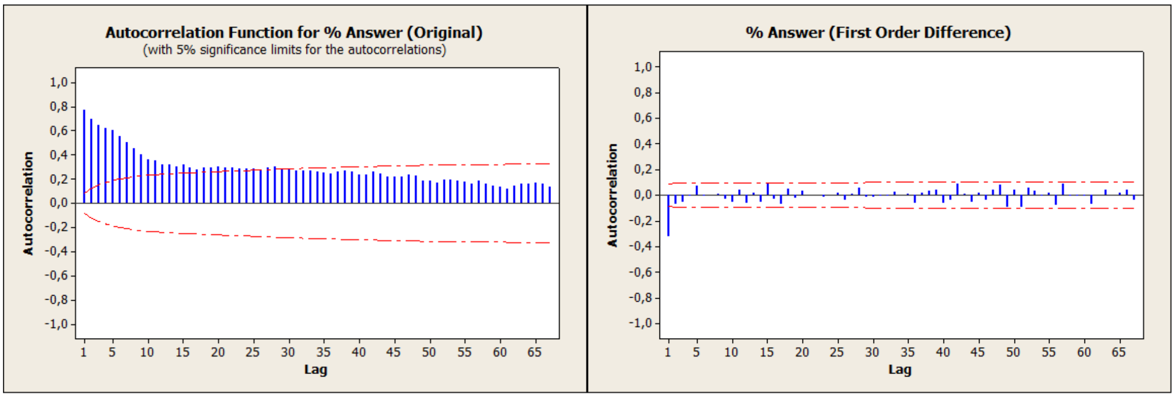

Figure B.1. Original and first order difference autocorrelation functions for *AP* (weekdays).

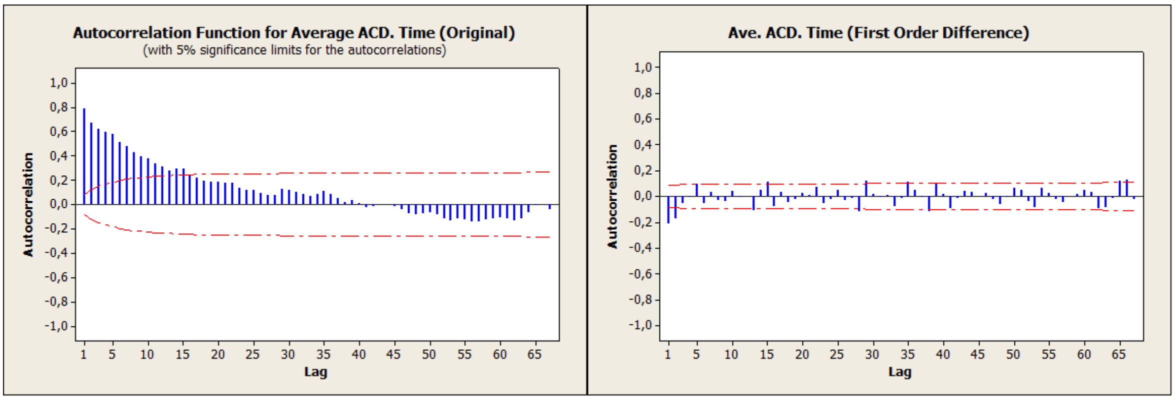

Figure B.2. Original and first order difference autocorrelation functions for $\overline{ACD}$ (weekdays).

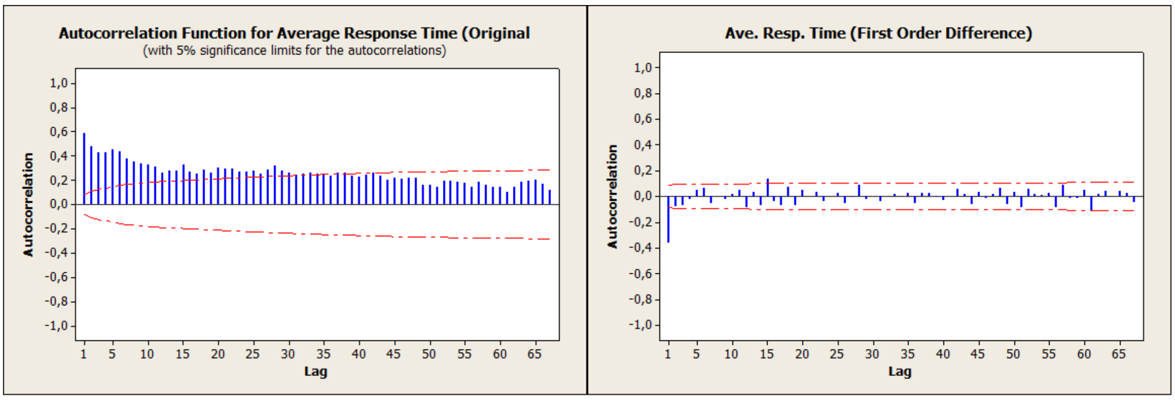

Figure B.3. Original and first order difference autocorrelation functions for $\overline{RT}$ (weekdays).

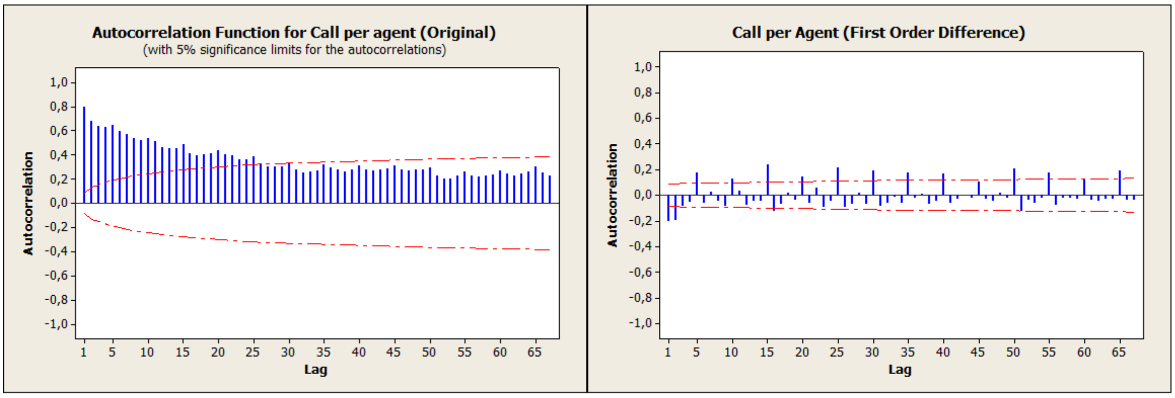

Figure B.4. Original and first order difference autocorrelation functions for *CpA* (weekdays).

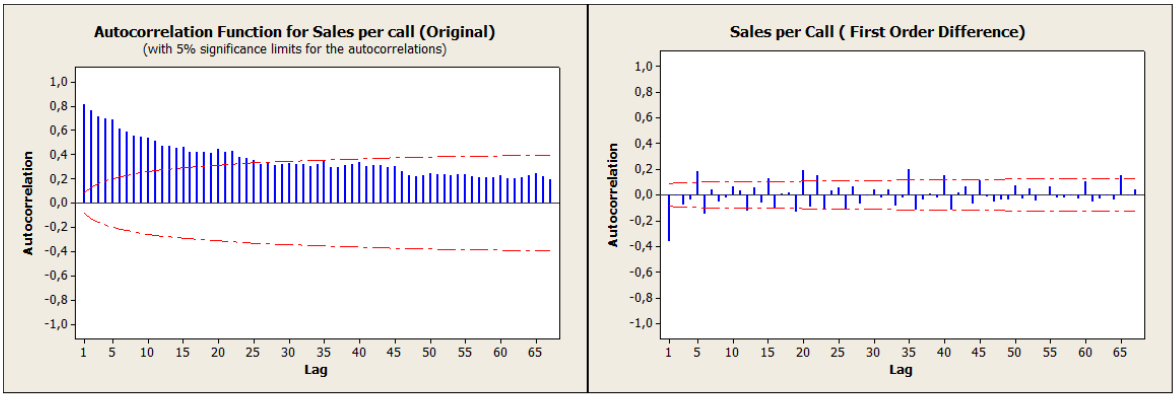

Figure B.5. Original and first order difference autocorrelation functions for *SpC* (weekdays).

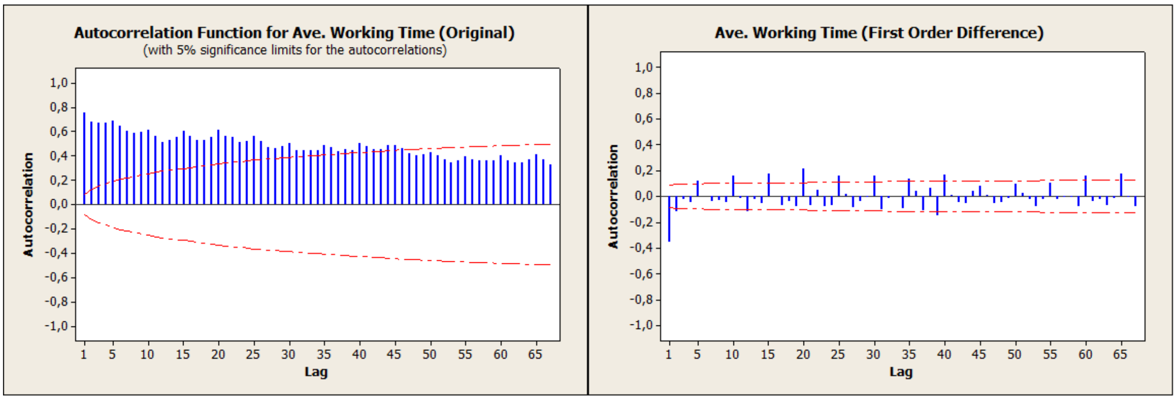

Figure B.6. Original and first order difference autocorrelation functions for $\overline{WT}$ (weekdays).

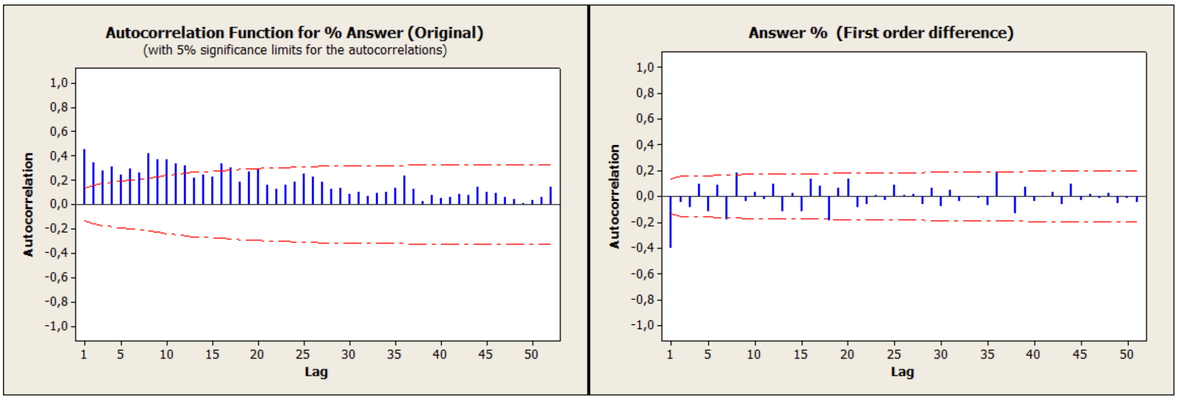

Figure B.7. Original and first order difference autocorrelation functions for *AP* (weekends).

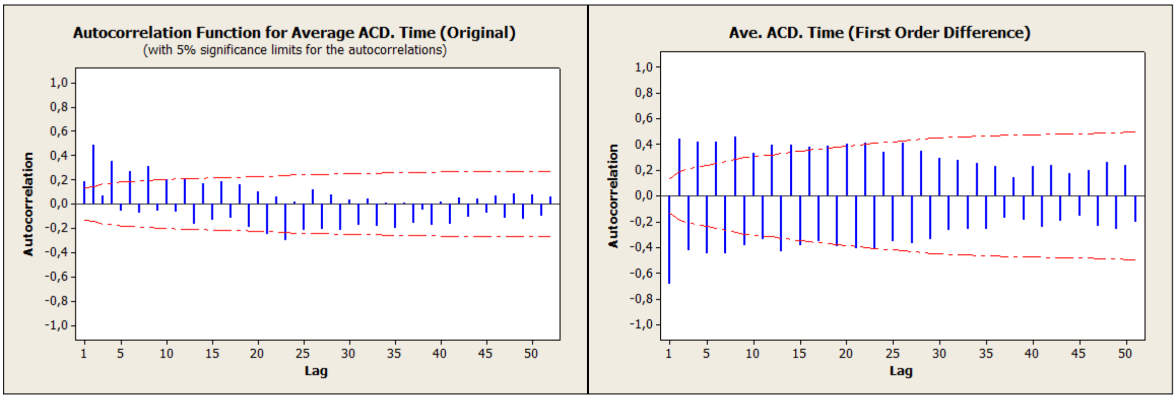

Figure B.8. Original and first order difference autocorrelation functions for $\overline{ACD}$ (weekends).

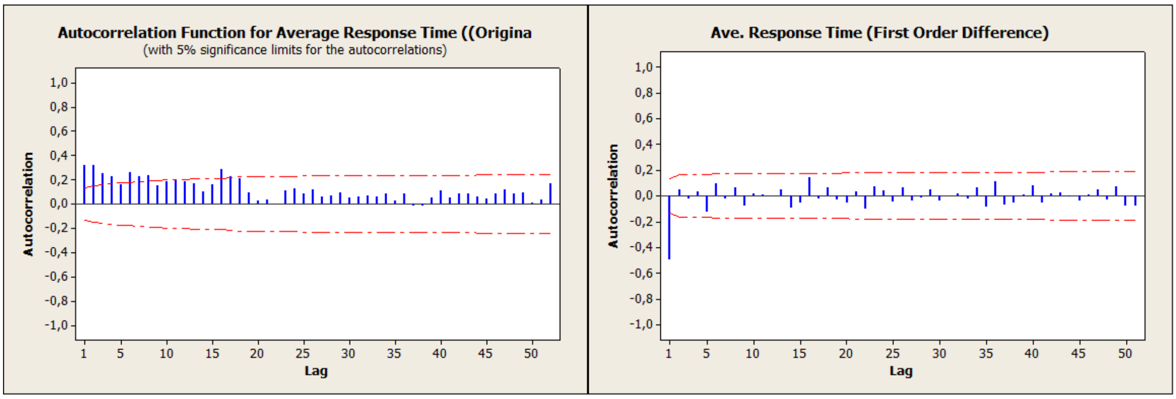

Figure B.9. Original and first order difference autocorrelation functions for $\overline{RT}$ (weekends).

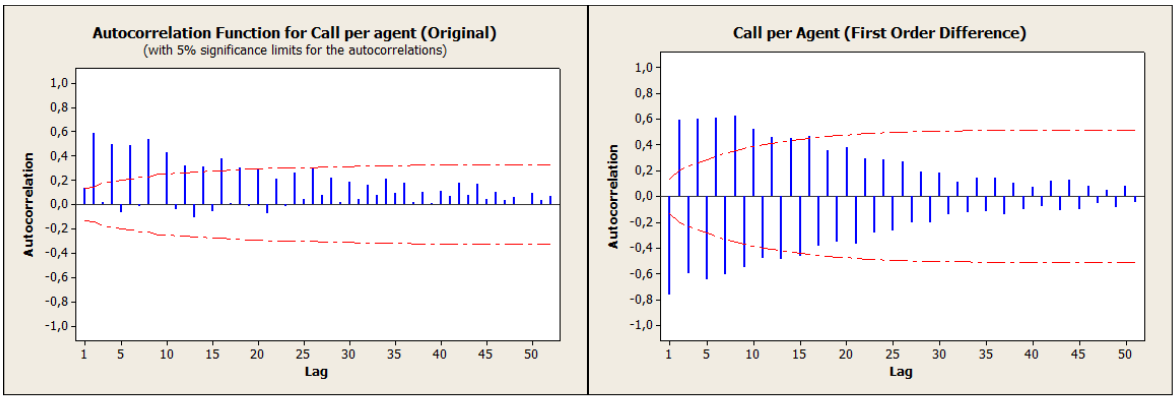

Figure B.10. Original and first order difference autocorrelation functions for *CpA* (weekends).

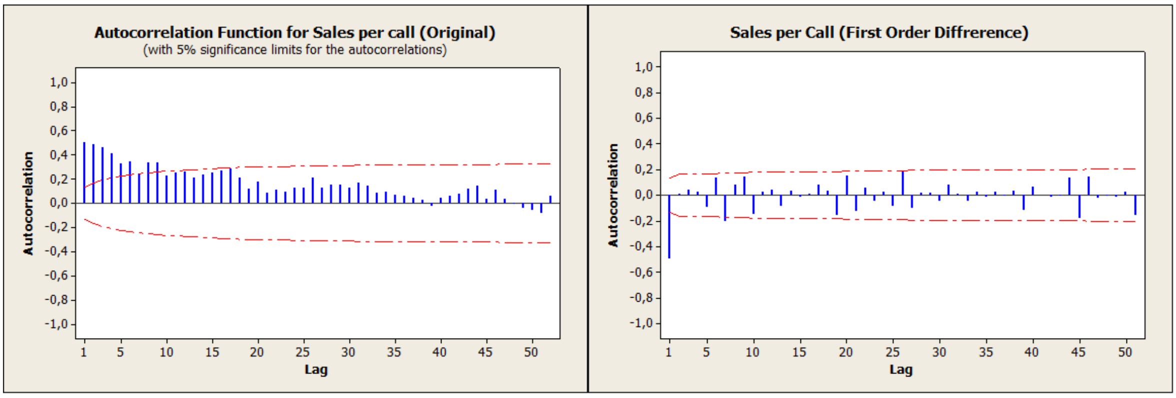

Figure B.11. Original and first order difference autocorrelation functions for *SpC* (weekends).

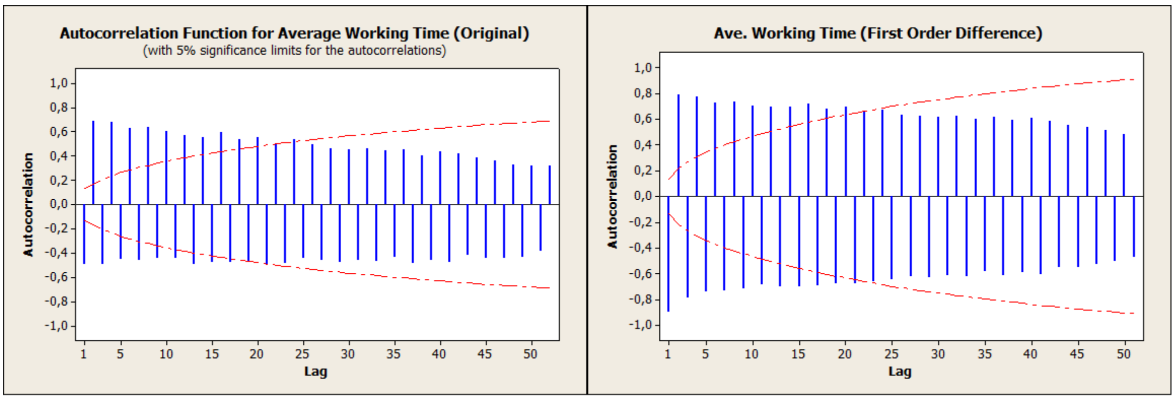

Figure B.12. Original and first order difference autocorrelation functions for $\overline{WT}$ (weekends).

APPENDIX C: AUTOCORRELATION FUNCTIONS OF WEEKDAYS AND WEEKENDS (DESEASONALIZED DATA)

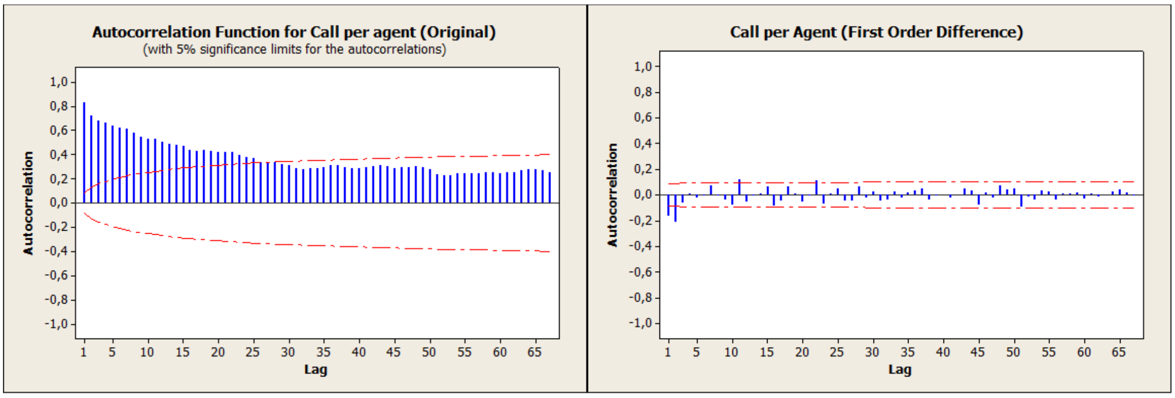

Figure C.1. Original and first order difference autocorrelation functions for deseasonalized *CpA* (weekdays).

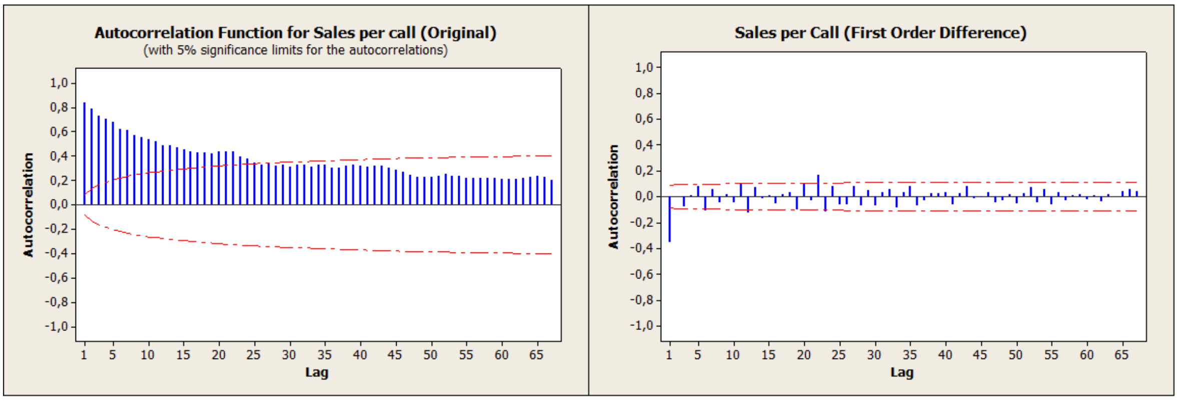

Figure C.2. Original and first order difference autocorrelation functions for deseasonalized *SpC* (weekdays).

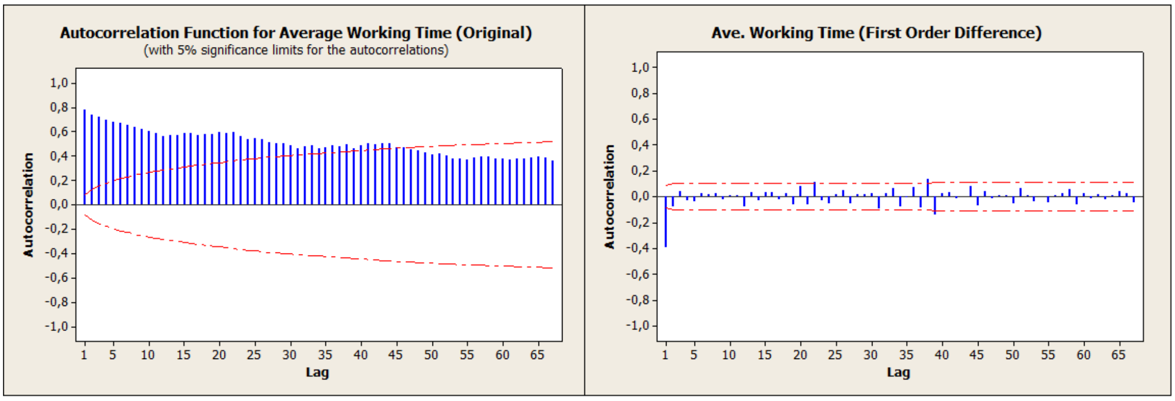

Figure C.3. Original and first order difference autocorrelation functions for deseasonalized $\overline{WT}$ (weekdays).

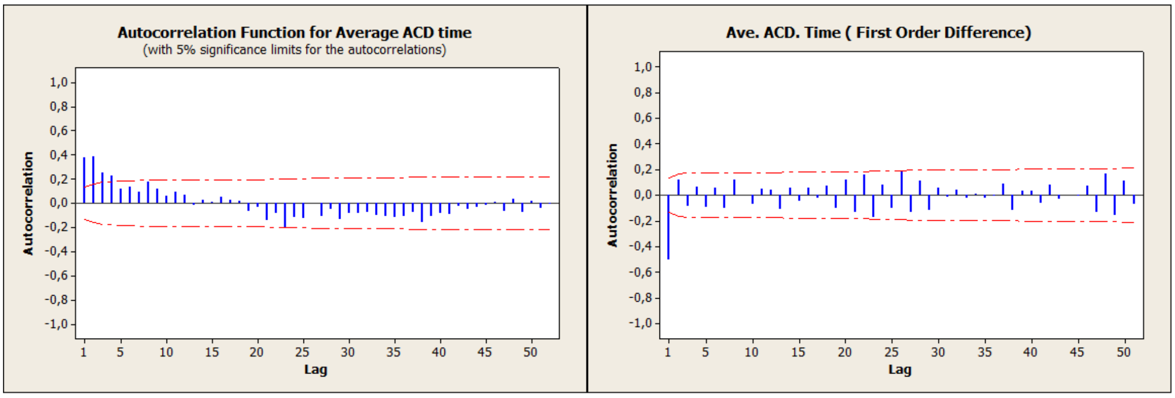

Figure C.4. Original and first order difference autocorrelation functions for deseasonalized $\overline{ACD}$ (weekends).

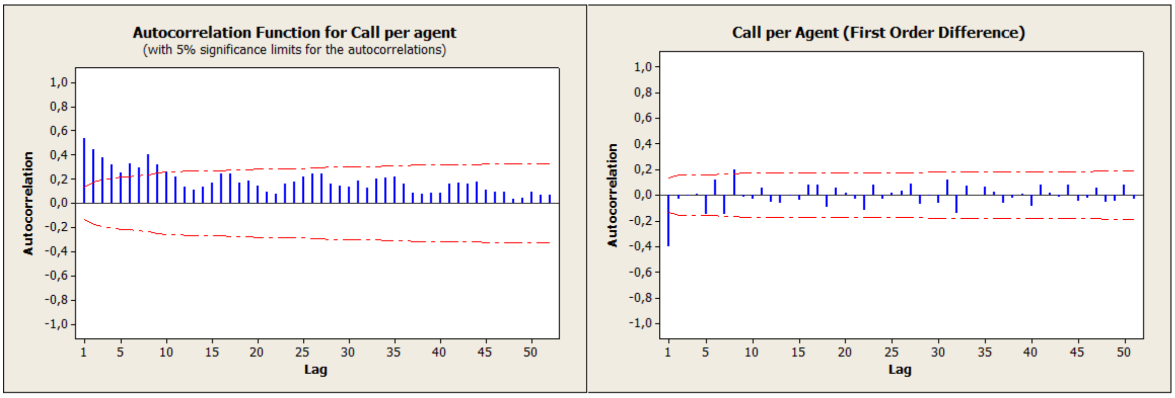

Figure C.5. Original and first order difference autocorrelation functions for deseasonalized CpA (weekends).

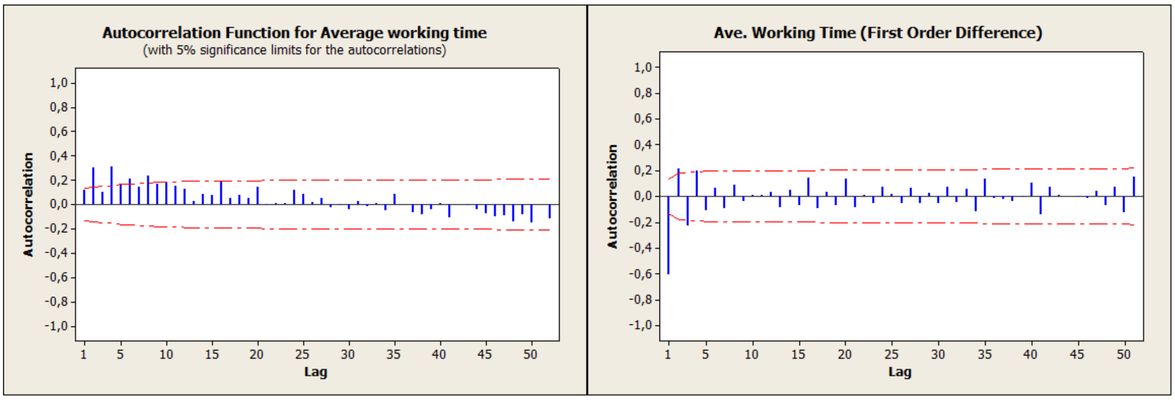

Figure C.6. Original and first order difference autocorrelation functions for deseasonalized $\overline{WT}$ (weekends).

APPENDIX D: MATLAB COMMANDS FOR THE PERIODIC WEIGHTED MOVING AVERAGE PCA ALGORITHM

```
%The input variables "X" "xt1" "period" "WF" should be given
%to the function.

%The "X" represents the HDS. It should be a matrix. The rows
%of "X" should
%contain the days, and the columns should contain the
%performance metrics

%The "xt1" represents the days to be evaluated by the
%algorithm. It should be a matrix.
%The rows of "xt1" should contain the performance metrics,
%and the columns
%should contain the days.

%The "period" represents the DT. It should be an integer
%which is smaller
%than the number of rows of the "X"

%The "WF" represents the weighting factor. It should be a
%number between 0
%and 1


function
[T2values,T2limits,T2warnings,T2signals,SPEvalues,SPElimits,S
PEwarnings,SPEsignals,Dstatistics,SPEcontrres,weightedmean,we
ightedstd,pcanumbers,counter,signals,warnings]              =
PWMAPCA(X,xt1,period,WF)


Xbart0=(1-WF)*mean(X(1:end-period-1,:))+(WF)*mean(X(end-
period:end,:));% Weighted mean
Xstddevt0=(1-WF)*std(X(1:end-period-1,:))+(WF)*std(X(end-
period:end,:));% Weighted std. dev.
Xstdt0=(X-
repmat(Xbart0,size(X,1),1))./repmat(Xstddevt0,size(X,1),1);%
%Standardize the HDS based on
% the weighted mean and weighted std. dev.

z=1;
k=1;
f=1;
counter=0;
```

```
while  k<=size(xt1,2)%  This  "while"  loop  works  for  all  the
%days that will be assessed through the model

weightedmean(:,k)=Xbart0;
weightedstd(:,k)=Xstddevt0;

xstdt1=(xt1(:,k)-Xbart0')./Xstddevt0';%standardize        the
%observation vector

addedstd=[Xstdt0;xstdt1'];%The temporary data matrix

covupdated0=(1-WF)*cov(addedstd(1:end-period-
1,:))+(WF)*cov(addedstd(end-period:end,:));%computes      the
%weighted covariance matrix

[COEFF,latent,explained]  =  pcacov(covupdated0);%  This  lines
%computes  the  eigenvectors  (principal  components),  eigen
%values and related percent variance explained values.

i=1;
top=explained(i,1);
while top<90
top=top+explained(i+1,1);
i=i+1;
end  %  determines  the  number  of  principal  components  to  be
%retained in the model=i

pcanumber=i;

pcanumbers(k,:)=pcanumber;

L=diag(latent);
W=COEFF/sqrt(L);
V=COEFF*sqrt(L);
xstdt1yscores=W(:,1:pcanumber)'*xstdt1;
     %  Next  2  command  lines  include  the  alternative  way  to
%compute
     %   the  T2  value  based  on  the  COEFF  matrix  (COEFF=U
%matrix) rather than W matrix.
%zzz=COEFF(:,1:pcanumber)'*xstdt1;
%T2zzz=zzz'*inv(L(1:pcanumber,1:pcanumber))*zzz;
T2=diag(xstdt1yscores'*xstdt1yscores);

% This "for" loop computes the "D" statistic
for g=1:size(xstdt1,1)
    Wexcl=W([1:g-1 g+1:end],1:pcanumber);
    yscoreexcluded=Wexcl'*xstdt1([1:g-1 g+1:end],:);
    T2excluded=diag(yscoreexcluded'*yscoreexcluded);
```

```
     % Next 3 command lines includes the alternative way to
%compute the D
     % statistic based on the "z" vector as stated in
%Equation 2.2 in the
     % thesis text.
  %zzzexcluded=COEFF([1:g-1
g+1:end],1:pcanumber)'*xstdt1([1:g-1 g+1:end],:);

%T2excludedzzz=zzzexcluded'*inv(L(1:pcanumber,1:pcanumber))*z
zzexcluded;
  %difference=(T2-T2excludedzzz);
    difference=(T2-T2excluded);
    Dstatistics(g,k)=difference;
end


% Next 3 command lines includes the computation of the SPE
%statistic
xstdt1sapka=V(:,1:pcanumber)*xstdt1yscores;
resxstdt1=xstdt1-xstdt1sapka;
SPE=diag(resxstdt1'*resxstdt1);

% Next 2 command lines includes the computation of the
%control and warning
%limits for the T2 statistic
T2limit=pcanumber*(size(X,1)-1)*(size(X,1)+1)*finv(1-
0.01,pcanumber,size(X,1)-pcanumber)/((size(X,1)-
pcanumber)*size(X,1));
T2warning=pcanumber*(size(X,1)-1)*(size(X,1)+1)*finv(1-
0.05,pcanumber,size(X,1)-pcanumber)/((size(X,1)-
pcanumber)*size(X,1));

%Next 14 command lines includes the compuatation of the
%control and warning
%limits for the SPE statistic
Lteta=L(pcanumber+1:size(X,2),pcanumber+1:size(X,2));
teta1=sum(diag(Lteta));
teta2=sum(diag(Lteta^2));
teta3=sum(diag(Lteta^3));
h0=1-(2*teta1*teta3)/(3*teta2^2);
if h0>0
calfa= norminv(0.99,0,1);
calfawarning=norminv(0.95,0,1);
else
    calfa= norminv(0.01,0,1);
calfawarning=norminv(0.05,0,1);
end
SPElimit=teta1*((calfa*sqrt(2*teta2*h0^2)/teta1)+(teta2*h0*(h
0-1)/teta1^2)+1)^(1/h0);
```

```
SPEwarning=teta1*((calfawarning*sqrt(2*teta2*h0^2)/teta1)+(te
ta2*h0*(h0-1)/teta1^2)+1)^(1/h0);


SPEcontrres(:,k)=resxstdt1;%   Contributions   for   the   SPE
statistic



T2values(k,:)=T2;
SPEvalues(k,:)=SPE;
T2limits(k,:)=T2limit;
T2warnings(k,:)=T2warning;
SPElimits(k,:)=SPElimit;
SPEwarnings(k,:)=SPEwarning;

if   T2limit>T2>T2warning   ||   SPElimit>SPE>SPEwarning   &&
T2<T2limit && SPE<SPElimit
    warnings(f,1)=k;
    f=f+1;
end

if T2<T2limit && SPE<SPElimit
    counter=counter+1;% Records the number of in-control days

X=[X;xt1(:,k)'];%Augments the HDS (i.e., it adds the current
%in-control day into the HDS)

Xbart0=(1-WF)*mean(X(1:end-period-1,:))+(WF)*mean(X(end-
period:end,:)); %Calculates the new weighted mean

Xstddevt0=(1-WF)*std(X(1:end-period-1,:))+(WF)*std(X(end-
period:end,:));%Calculates the new weighted std. dev.

Xstdt0=(X-
repmat(Xbart0,size(X,1),1))./repmat(Xstddevt0,size(X,1),1);%
%Standardize the new HDS based on
% the new weighted mean and new weighted std. dev.

else

signals(z,1)=k;%  adds  the  out-of-control  day  into  signals
%vector
z=z+1;
end
k=k+1;
end
l=1;
for w=1:size(T2values,1)
    if T2values(w,1)>T2limits(w,1)
```

```
        T2signals(l,1)=w;
        l=l+1;
    end
end

h=1;
for w=1:size(SPEvalues,1)
    if SPEvalues(w,1)>SPElimits(w,1)
        SPEsignals(h,1)=w;
        h=h+1;
    end
end

end
```

APPENDIX E: MATLAB COMMANDS FOR THE MONTE CARLO STUDY (FIRST POLICY)

```
% "C" "mu" "hdslength" "signalnumber" "DP" "WF" variable
%should be given
% to the montecarlorand function as the input varaiables.

% "C" is p x p lower triangular matrix produced by Cholesky
decomposition
%of the intended covariance matrix of the data to be
generated.

% "mu" is pxn mean vector.

% "hdslenght" is a positive integer. It represents the number
%of the days which will be taken as the
%HDS.

% "signalnumber" is a positive integer. It represents the
%number of variables that will be manipulated within a
%particular day. Itshould be lower than or equal to p


% "DP" is a positive integer. It represents the number of
%days that will be
%included in the dominant period.

% " WF" is a positive integer. It represents the value of the
%weighting factor.


function [meanpower,unnecessarypercent] =
montecarlorand(C,mu,hdslength,signalnumber,DP,WF)

for i=1:100 % this "for" loop executes 100 iterations
Z=randn(size(mu,1),size(mu,2));
X=mu+C*Z; % generates the multinormal data
mu100=mu(:,hdslength+1:end);
mut=mu100';
Xt=X';
hds=Xt(1:hdslength,:); % selects the HDS
t=Xt(hdslength+1:end,:); % assign the remaining days as days
%to be evaluated
sigma=std(Xt);

% Next 19 lines injects the signals in randomly chosen days.
%The number of variables to be manipulated within each day is
%determined by "signalnumber"
```

```
ensr=1:100;
for q=1:10
    indexr=randint(1,1,[1,size(ensr,2)]);
    r=ensr(1,indexr);
    rvector(q,1)=r;% records the randomly chosen 10 days in
%which the signals are injected
    ensr=ensr(1,[1:indexr-1 indexr+1:end]);
    ensc=1:6;
    for s=1:signalnumber
    indexc=randint(1,1,[1,size(ensc,2)]);
    c=ensc(1,indexc);
    cvector(q,s)=c;% records the randomly chosen manipulated
%variables within each day
    ensc=ensc(1,[1:indexc-1 indexc+1:end]);
    if randint==0
t(r,c)=mut(r,c)-3*sigma(1,c);%manipulate the variable
    else
t(r,c)=mut(r,c)+3*sigma(1,c);%manipulate the variable
    end
    end
end

% From this line on the generated data set is evaluated via
%the proposed PWMAPCA algorithm.
[T2values,T2limits,T2warnings,T2signals,SPEvalues,SPElimits,S
PEwarnings,Dstatistics,SPEcontrres,weightedmean,weightedstd,p
canumbers,counter,signals] = PWMAPCA(hds,t',DP,WF);
p=0;
u=0;
for k=1:size(signals,1)

  if                     signals(k,1)==rvector(1,1)          ||
signals(k,1)==rvector(2,1)  ||  signals(k,1)==rvector(3,1)  ||
signals(k,1)==rvector(4,1)    ||    signals(k,1)==rvector(5,1)
||signals(k,1)==rvector(6,1) || signals(k,1)==rvector(7,1)  ||
signals(k,1)==rvector(8,1)  ||  signals(k,1)==rvector(9,1)  ||
signals(k,1)==rvector(10,1)
      p=p+1;
  else
      u=u+1;
  end
end
    unnecessarsignals(i,1)=u/size(X,2);     %  Computes  the
%Pr{unwanted signal} for a particular iteration

  power=p/10;% Computes the Pr{detection of injected signals}
%for a particular iteration

  powervector(i,1)=power;  %records  the  Pr{detection  of
injected signals} for each iteration
```

```
end
unnecessarypercent=mean(unnecessarsignals);  %Calculates  the
mean  Pr{unwanted signal} for 100 iterations
meanpower=mean(powervector);%Calculates the mean Pr{detection
of injected signals} for 100 iterations
end
```

APPENDIX F: MATLAB COMMANDS FOR THE MONTE CARLO STUDY (SECOND POLICY)

```
% "C" "mu" "hdslength"  "DP" "WF" variables should be given
% to the montecarlo function as the input varaiables.

% "C" is p x p lower triangular matrix produced by Cholesky
%decomposition of the intended covariance matrix of the data
%to be generated.

% "mu" is pxn mean vector.

% "hdslenght" is a positive integer. It represents the number
%of the days which will be taken as the HDS.

% "DP" is a positive integer. It represents the number of
%days that will be included in the dominant period.

% " WF" is a positive integer. It represents the value of the
% weighting factor.

function [meanpower,unnecessarypercent] =
montecarlo(C,mu,hdslength,DP,WF)
for i=1:100% this "for" loop executes 100 iterations
Z=randn(size(mu,1),size(mu,2));
X=mu+C*Z;% generates the multinormal data
mu100=mu(:,hdslength+1:end);
mut=mu100';
Xt=X';
hds=Xt(1:hdslength,:);% selects the HDS
t=Xt(hdslength+1:end,:);% assign the remaining days as days
%to be evaluated
sigma=std(Xt);

% 5,15,25,40,45,60,70,80,95,100  are the arbitrarily chosen
%days in which the signals are injected.
%The columns of t matrix represents which variables are
manipulated.
%In one-variable signal case the %variables with 1 "%" are
manipulated.
%In two-variable signal case the %variables with 1 and 1 "%"
were manipulated.
%In three-variable signal case all three variables were
manipulated.

t(5,1)=mut(5,1)-3*sigma(1,1);%%
t(5,3)=mut(5,3)+3*sigma(1,3);%
t(5,4)=mut(5,4)+3*sigma(1,4);%%%
```

```
t(15,2)=mut(15,2)-3*sigma(1,2);%%
t(15,4)=mut(15,4)+3*sigma(1,4);%%%
t(15,5)=mut(15,5)-3*sigma(1,5);%

t(25,2)=mut(25,2)-3*sigma(1,2);%
t(25,5)=mut(25,5)+3*sigma(1,5);%%
t(25,6)=mut(25,6)+3*sigma(1,6);%%%

t(40,6)=mut(40,6)+3*sigma(1,6);%
t(40,4)=mut(40,4)-3*sigma(1,4);%%
t(40,3)=mut(40,3)-3*sigma(1,3);%%%

t(45,2)=mut(45,2)+3*sigma(1,2);%%%
t(45,5)=mut(45,5)+3*sigma(1,5);%%
t(45,3)=mut(45,3)+3*sigma(1,3);%

t(60,4)=mut(60,4)+3*sigma(1,4);%
t(60,3)=mut(60,3)+3*sigma(1,3);%%
t(60,5)=mut(60,5)-3*sigma(1,5);%%%

t(70,6)=mut(70,6)+3*sigma(1,6);%%
t(70,5)=mut(70,5)-3*sigma(1,5);%
t(70,1)=mut(70,1)-3*sigma(1,1);%%%

t(80,6)=mut(80,6)-3*sigma(1,6);%
t(80,4)=mut(80,4)+3*sigma(1,4);%%
t(80,1)=mut(80,1)-3*sigma(1,1);%%%

t(95,1)=mut(95,1)-3*sigma(1,1);%%
t(95,3)=mut(95,3)+3*sigma(1,3);%
t(95,2)=mut(95,2)+3*sigma(1,2);%%%

t(100,3)=mut(100,3)+3*sigma(1,3);%%
t(100,5)=mut(100,5)+3*sigma(1,5);%%%
t(100,6)=mut(100,6)+3*sigma(1,6);%


% From this line on the generated data set is evaluated via
%the proposed PWMAPCA algorithm.

[T2values,T2limits,T2warnings,T2signals,SPEvalues,SPElimits,S
PEwarnings,Dstatistics,SPEcontrres,weightedmean,weightedstd,p
canumbers,counter,signals] = PWMAPCA(hds,t',DP,WF);
p=0;
u=0;
for k=1:size(signals,1)

  if  signals(k,1)==5 || signals(k,1)==15 || signals(k,1)==25
|| signals(k,1)==40 || signals(k,1)==45 ||signals(k,1)==60 ||
```

```
signals(k,1)==70 || signals(k,1)==80 || signals(k,1)==95 ||
signals(k,1)==100
      p=p+1;
  else
      u=u+1;
  end
end
    unnecessarsignals(i,1)=u/size(X,2);     % Computes the
Pr{unwanted signal} for a particular iteration

  power=p/10;% Computes the Pr{detection of injected signals}
for a particular iteration

  powervector(i,1)=power;%records the Pr{detection of
injected signals} for each iteration
end
unnecessarypercent=mean(unnecessarsignals);%Calculates the
%mean  Pr{unwanted signal} for 100 iterations
meanpower=mean(powervector);%Calculates the mean Pr{detection
of injected signals} for 100 iterations
end
```

REFERENCES

Alfaro, E., J. L. Alfaro, M. Gamez, and N. Garcia, 2009, “A Boosting Approach for Understanding out-of-control Signals in Multivariate Control Charts”, *International Journal of Production Research* , Vol. 47, No. 24, pp. 6821-6834.

AlGhazzawi, A., and B. Lennox, 2009, “Model Predictive Control Using Multivariate Statistics”, *Journal of Process Control* , Vol. 19, No. 2, pp. 314-327.

AlGhazzawi, A., and B. Lennox, 2008, “Monitoring a Complex Refining Process Using Multivariate Statistics”, *Control Engineering Practice* , Vol. 16, No. 3, pp. 294-307.

Box, G. E., and G. M. Jenkins, 1970, *Time Series Analysis,* Holden Day Inc., California.

Call Centers Association, 2010, *Dünya ve Türkiye Çağrı Merkezleri Sektörü,* http://cagrimerkezleridernegi.org/index.php, accessed at February 2011.

Chang, Y. S., 2007, “Multivariate CUSUM and EWMA Control Charts for Skewed Populations Using Weighted Standart Deviations”, *Communications in Statistics—Simulation and Computation* , Vol. 36, pp. 921-936.

Cleveland, B., 2007 , “The Measures Every Successful Call Center Should Have -- There are seven types of measures that should be in place in every customer contact center”, *Call Center Magazine,* Vol. 20, No. 4, p. 6.

Corbett, C. J., and J. N. Pan, 2002 , “Evaluating Environmental Performance Using Statistical Process Control Techniques”, *European Journal of Operational Research,* Vol. 139, No. 1, pp. 68-83.

Croiser, R. B., 1988, “Multivariate Generations of Cumulative Sum Quality Control Schemes”, *Technometrics,* Vol. 30, pp. 219-303.

Dabholkar, P., C. Shepherd, and D. Thorpe, 2000, “A Comprehensive Framework for Service Quality: An Investigation of Critical Conceptual and Measurement Issues Through a Longitudinal Study”, *Journal of Retailing* , Vol.18, No. 2, pp. 139-173.

Davier, M., S. Sinharay, O. Andreas, and A. Beaton, 2006, “The Statistical Procedures Used in National Assessment of Educational Progress: Recent Developments and Future Directions”, *Handbook of Statistics,* Vol. 26, pp. 1039-1055.

Economic Policy Research Foundation of Turkey, 2011, *Turkcell Global Bilgi Erzurum Çağrı Merkezi Ekonomik Etki Değerlendirme Çalışması,* http://www.tepav.org.tr/, accessed at September 2011.

Elaine, J. L., 2010, “Improving Customer Retention Through Service Quality in Call Centers”, *International Journal of Management and Infırmation Systems,* Vol. 14, No. 3, p. 71.

Esposito, M., and M. Tenenhaus, 2008, “Statistical Methods In Performance analysis”, *Applied Stochastic Models in Business and Industry,* Vol. 24, pp. 369-371.

Floss, D., 2007, “Contact Center Performance Management Guide -- What is it? What does it accomplish? And should you be using it?”, *Call Center Magazine,* Vol. 20, No. 3, p. 12.

Funch, C., and R. S. Kenett, 1998, *Multivariate Quality Control Theory and Applications,* Marcel Dekker Inc., New York.

Gans, N., G. Koole, and A. Mandelbaum, 2003, “Telephone Call Centers: Tutorial, Review and Research Prospects”, *Manufacturing and Service Operations Management,* Vol. 5, No. 2, pp. 79-141.

Geladi, P., and B. R. Kowalski, 1986, “Partial Least Squares Regression: A Tutorial”, *Analytica Chimica Acta,* Vol. 185, pp. 1-17.

He, S. G., L. Li, and E. S. Qi, 2007, "Study on the Continuous Quality Improvement of Telecommunication Call", *International Conference on Service Systems and Service Management,* Changdu - China, 2007, Institute of Electrical and Electronical Engineers Computer Society, United States.

Hotelling, H., 1947, "Multivariate Quality Control", In: C. Eisenhart, M. W. Hastay, and W. A. Wallis, *Techniques of Statistical Analysis*, New-York, 1947, McGraw-Hill, New-York.

Huo, J., and T. Dai, 2008, "Performance Evaluation of Multi-skill Call Center Considering Call Abandonment", *4th International Conference on Wireless Communications, Networking and Mobile Computing*, Dalian - China, 2008, IEEE, United States.

Jackson, J. E., 1991, *A User's Guide to Principal Components,* John Wiley and Sons Inc., New York.

Jaiswal, A. K., 2008, "Customer Satisfaction and Service Quality Measurement in Indian Call Centers", *Managing Service Quality,* Vol.18, No.4, pp. 405-416.

Jiang, W., T. Au, and K. L. Tsui, 2007, "A Statistical Process Control Approach to Business Activity Monitoring", *IIE Transactions,* Vol. 39, No. 3, pp. 235-249.

Jing, L., and G. Min, 2010, "Predicting Call Center Service Grade with Improved Neural Network Algorithm", *2nd International Workshop on Intelligent Systems and Applications,* Wuhan - China, 2010, IEEE, United States.

Kornyshev, Y. N., and V. I. Duz, 1990, "Accuracy of Evaluating Call Servicing Quality Characteristics in Communication Networks", *Telecommunications and Radio Engineering,* Vol. 45, No. 4, pp. 1-8.

Lane, S., E. B. Martin, A. J. Morris, and P. Gower, 2003, "Application of Exponentially Weighted Principal Component Analysis for the Monitoring of a Polymer Film

Manufacturing Process", *Transactions of the Institute of Measurement and Control,* Vol. 25, pp. 17-35.

Leu, S. S., and Y. C. Lin, 2008, "Project Performance Evaluation Based on Statistical Process Control Techniques", *Journal of Construction Engineering and Management,* Vol. 134, No. 10, pp. 813-819.

Li, W., H. H. Yue, S. V. Cervantes, and S. J. Qin, 2000, "Recursive PCA for Adaptive Process Monitoring", *Journal of Process Control,* Vol. 10, pp. 471-486.

Li, Y., Y. Cheng, L. Zhang, and H. Huang, 2008, "Application of Multivariate Statistical Analysis to Classify Electricity Customers", *2008 China International Conference on Electricity Distribution,* Guangzhou - China, 2008, IEEE Computer Society, United States.

Lowry, C. A., W. H. Woodall, C. W. Champ, and S. E. Rigdon, 1992, "A Multivariate Exponentially Weighted Moving Average Control Chart", *Technometrics,* Vol. 34, No. 1, pp. 46-53.

Lu, C. W., and M. R. Reynolds, 2001, "Cusum Control Charts for Monitoring an Autocorrelated Process", *Journal of Quality Technology,* Vol. 33, No. 3, pp. 316-334.

Lu, C. W., and M. R. Reynolds, 1999, "EWMA Control Charts for Monitoring the Mean of Autocorrelated Process", *Journal of Quality Technology,* Vol. 31, No. 2, pp.166-188.

Ma, J., N. Kim, and L. Rothrock, 2011, "Performance Assessment in an Interactive Call Center Workforce Simulation", *Simulation Modelling Practice and Theory,* Vol. 19, pp. 227-238.

Mason, R. L., and J. C. Young, 2002, *Multivariate Statistical Process Control with Industrial Applications,* American Statistical Association and Society for Industrial and Applied Mathematics, Philadelphia.

Mason, R. L., N. D. Tracy, and J. C. Young, 1995, “Decomposition of T^2 for Multivariate Control Chart Interpretation”, *Journal of Quality Technology*, Vol. 27, No. 2, pp. 99-108.

Mastrangelo, C. M., G. C. Runger, and D. C. Montgomery, 1996, “Statistical Process Monitoring with Principal Components”, *Quality and Reliability Engineering International,* Vol. 12, pp. 203-210.

Mehrotra, V., and J. Fama, 2003, “Call Center Simulation Modeling: Methods, Challenges and Opportunities”, *Proceedings of the 2003 Winter Simulation Conference, New Orleans - USA, 2003, IEEE, United States.*

Miletic, I., S. Quinn, M. Dudzic, V. Vaculik, and M. Champagne, 2004, “An Industrial Perspective on Implementing on-line Applications of Multivariate Statistics”, *Journal of Process Control* , Vol. 14, No. 8, pp. 821-836.

Ming, Z., J. Yong, S. Xinming, and T. Kuo, 2009, “Risk Evaluation of Power Supplying Enterprises Based on PCA and BP Neural Network”, *International Conference on Management and Service Science,* Wuhan - China, 2009, IEEE Computer Society, United States.

Mujica, L. E., J. Vehi, M. Ruiz, M. Verleysen, W. Staszewski, and K. Worden, 2008, “Multivariate Statistics Process Control for Dimensionality Reduction in Structural Assessment”, *Mechanical Systems and Signal Processing,* Vol. 22, No. 1, pp. 155-171.

Paprzycki, M., A. Abraham, R. Guo, and S. Mukkamala, 2004, “Data Mining Approach for Analyzing Call Center Performance”, *17th International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert Systems,* Ottawa - Canada, 2004, Springer Verlag, Germany.

Pearson, R., 2002, “Outliers in Process Modelling and Identification”, *IEEE Transactions on Control Systems Technology,* Vol. 10, No. 1, pp. 55-63.

Purcell, J. A., 1994, “Statistical Process Control Charting Applied to the Analysis of Human Performance in Computer-Supported Environments”, *Proceedings of 38th Annual Meeting of the Human Factors and Ergonomics Society,* Nashville - USA, 1994, Human Factors & Ergonomics Soc., United States.

Ramseook-Munhurrun, P., P. Naidoo, and S. D. Lukea-Bhiwajee, 2009, “Employee Perceptions of Service Quality in a Call Center”, *Managing Service Quality,* Vol. 19, No. 5, pp. 541-557.

Rex, S. G., 1999, “The Application of Statistical Process Control to Manage Global Client Outcomes in Behavioral Healthcare”, *Evaluation and Program Planning,* Vol. 22, No. 2, pp. 199-210.

Runger, G. C., 2002, “Assignable Causes and Autocorrelation: Control Charts for Observations or Residuals", *Journal of Quality and Technology,* Vol. 34, No. 2, pp. 165-170.

Runger, G. C., F. B. Alt, and D. C. Montgomery, 1996, “Contributors to a Multivariate Statistical Process Control Chart Signal”, *Communications in Statistics - Theory and Methods,* Vol. 25, pp. 2203-2213.

Sandoz, D. J., 2003, “The Exploitation of Adaptive Modelling in the Model Predictive Control Environment of Connoiseur”, In: V. VanDoren, *Techniques for Adaptive Control* , Butterworth-Heinemann, United States.

Serra, A. P., and M. Martucci Jr., 2003, “The Application of CRM in Call Centers”, *Management Information Systems,* Vol. 7, pp. 367-376.

Shehab, R. L., R. E. Schlegel, and G. Kirby, 1997, “The Use of Statistical Quality Control Charts to Evaluate Changes in Individual Performance”, *Proceedings of the 1997*

41st Annual Meeting of the Human Factors and Ergonomics Society, Albuquerque - USA, 1997, Human Factors and Ergonomics Society Inc., United States.

Surtihadi, J., M. Raghavachari, and G. Runger, 2004, "Multivariate Control Charts for Process Dispersion", *International Journal of Production Research,* Vol. 42, No. 15, pp. 2993–3009.

Tsai, M. C., and C. Ou-Yang, 2010, "Improving a Commercial Bank's Operation Performance Through Statistical Process Control", *Journal of the Chinese Institute of Industrial Engineers,* Vol. 27, No. 3, pp. 226-236.

Wang, Q., Jiang, N., Gou, L., Che, M., and Zhang, R., 2006. Practical Experiences of Cost/Schedule Measure Through Earned Value Management and Statistical Process Control. *International Software Process Workshop and International Workshop on Software Process Simulation and Modeling.* Springer Verlag.

Wang, X., U. Kruger, and B. Lennox, 2003, "Recursive Partial Least Square Algorithms for Monitoring Complex Industrial Processes", *Control Engineering Practice,* Vol. 11, No. 6, pp. 613-632.

Whitt, W., 2006, "The Impact of Increased Employee Retention on Performance in a Customer Contact Center", *Manufacturing and Service Operations Management,* Vol. 8, No. 3, pp. 235-252.

Wold, S., 1994, "Exponentially Weighted Moving Principal Components Analysis and Projection to Latent Structures", *Chemometrics and Intelligent Laboratory Systems,* Vol. 23, pp. 149-161.

Xiong, L., J. Liang, and J. Qian, 2006, "Performance Monitoring of Chemical Process Based on Multivariate Statistical Technology", *Sixth World Congress on Intelligent Control and Automation,* Dalian - China, 2006, IEEE, United States.

Zhu, R., and Y. Zhu, 2009, “Performance Analysis of Call Centers Based on M/M/s/k+G Queue with Retrial, Feedback and Impatience”, *2009 IEEE International Conference on Grey Systems and Intelligent Services*, Nanjing - China, 2009, IEEE, United States.